



INSTITUTO DE HIGIENE E
MEDICINA TROPICAL
DESDE 1902



UNIVERSIDADE
NOVA
DE LISBOA

Universidade Nova de Lisboa
Instituto de Higiene e Medicina Tropical

**Origem e filogeografia do VIH-1 no arquipélago de São
Tomé e Príncipe**

Elisângela Jessica Vera Cruz Bonfim

DISSERTAÇÃO PARA A OBTENÇÃO DO GRAU DE MESTRE EM SAÚDE PÚBLICA E
DESENVOLVIMENTO



INSTITUTO DE HIGIENE E
MEDICINA TROPICAL
DESDE 1902



UNIVERSIDADE
NOVA
DE LISBOA

Universidade Nova de Lisboa
Instituto de Higiene e Medicina Tropical

**Origem e filogeografia do VIH-1 no arquipélago de
São Tomé e Príncipe**

Autora: Elisângela Jessica Vera Cruz Bonfim

Orientadora: Dra Ana Abecasis (IHMT/UNL)

Coorientador: Dr. Victor Pimentel (IHMT/UNL)

Dissertação apresentada para cumprimento dos requisitos necessários à obtenção do grau de mestre em Saúde Pública e Desenvolvimento

MARCO, 2020

Toda a Escritura é inspirada por Deus, e útil para o ensino,
para a repreensão, para a correção e para a instrução na justiça,
para que o homem de Deus seja apto e plenamente
preparado para toda a boa obra.
(2 Timóteo 3:16-17)

Agradecimentos

Os sinceros agradecimentos vão em primeiro lugar à Deus, por ter me enchido de forças nos momentos mais difíceis, por me ter ensinado a batalhar e a lidar com as adversidades da vida; e por me mostrar o caminho da persistência, paciência e resiliência.

Não em segundo, mas os mais importantes na minha vida, os meus queridos pais Victor Manuel Sacramento Bonfim e Eulália Sebastiana Bandeira Vera Cruz, e minha irmã Vanda Solange Vera Cruz Bonfim, pelo apoio e amor incondicional durante todo o mestrado.

Aos meus primos, familiares, amigos e tios, nomeadamente João e Antónia Penetra, pela recepção e hospitalidade em sua casa.

Aos meus orientadores Dra. Ana Barroso Abecasis, e Dr. Victor Figueiredo Pimentel pela disponibilidade, paciência, orientações e sugestões. E também ao Dr. Bonifácio Sousa, diretor do PNLS (Programa Nacional de Luta contra a SIDA), pela disponibilização dos dados do programa.

E finalmente a todos aqueles que de alguma forma participaram e tiveram algum contributo nesta trajetória recheada de ensinamentos e amadurecimento.

Um bem-haja a todos.

Resumo

O VIH é um retrovírus que já causou a morte de cerca de 940 mil pessoas até 2017, e infetou mais de 36,9 milhões no mundo, sendo o continente africano a região do globo mais afetada pela epidemia. Este vírus altamente replicativo, recombinativo e com uma notável capacidade de mutação e adaptação às novas condições do ambiente humano, possui uma extraordinária diversidade genética global.

O VIH-1 é o tipo mais predominante e disperso a nível global, e o principal agente da SIDA. Este vírus é classificado em grupos, subtipos, sub-subtipos, formas recombinantes circulantes (CFRs) e formas únicas circulantes (URFs).

O grupo M do VIH-1 é o mais prevalente globalmente, e por isso estudar e conhecer a epidemiologia molecular deste grupo reveste-se de grande importância para a saúde mundial, para a melhoria da terapêutica ARV, para o alcance dos objetivos 90-90-90 definidos nos ODS (Objetivos do desenvolvimento sustentável) e para o desenvolvimento de vacinas.

S.Tomé e Príncipe, é um pequeno arquipélago localizado na costa ocidental africana, caracterizado por prevalência baixa (0,5%) do VIH-1, mas com poucos estudos realizados na área da epidemiologia molecular das estirpes VIH-1 circulantes no país.

Este trabalho de investigação tratou-se de um estudo observacional baseado na análise de sequências *pol* do VIH-1 colhidas em seropositivos, no período de 2004 a 2008 em S.Tomé e Príncipe. Com estas amostras foi possível determinar a diversidade genética das estirpes do VIH-1 que circulam no país, identificar as mutações de resistência aos antirretrovirais existentes e reconstruir a origem temporal e geográfica da epidemia a partir da estirpe mais prevalente identificada (CFR02_AG), através da reconstrução filogenética de uma árvore MCC (Maximum Credibility Clade) dessas sequências usando a estrutura filogeográfica bayesiana disponibilizadas no pacote do BEAST.

Palavras-chaves: VIH, S.Tomé e Príncipe, Epidemiologia molecular, Resistência aos ARVs

Abstract

HIV is a retrovirus that has killed about 940,000 people by 2017, and has infected more than 36.9 million worldwide, with Africa being the region most affected by the epidemic. This highly replicative, recombinant virus, with a remarkable capacity for mutation and adaptation to changing conditions of the human environment, has extraordinary global genetic diversity.

HIV-1 is the most prevalent and widespread type worldwide, and the main agent of AIDS. This virus is classified into groups, subtypes, subtypes, circulating recombinant forms (CFRs) and unique recombinant forms (URFs).

HIV-1 group M is the most prevalent globally, so studying and knowing its molecular epidemiology has a great importance for global health, for the improvement of ARV therapy, for the achievement of objectives 90-90- 90 defined in the SDGs (Sustainable Development Goals) and for vaccine development.

S. Tomé and Príncipe is a small archipelago located on the West African coast, characterized by low prevalence (0.5%) of HIV-1, but with few studies in the molecular epidemiology of HIV-1 strains circulating in the country.

This research was an observational study with HIV-1 pol sequences collected from HIV-positive people from 2004 to 2008 in S. Tomé and Príncipe. With these samples it was possible to determine the genetic diversity of HIV-1 strains circulating in the country, to identify existing antiretroviral resistance mutations and to reconstruct the temporal and geographical origin of the epidemic from the most prevalent strain identified (CFR02_AG) by phylogenetic reconstruction of a Maximum Credibility Clade (MCC) tree of these sequences using the Bayesian phylogeographic framework provided in the BEAST package.

KEYWORDS: HIV, S. Tomé and Príncipe, Molecular Epidemiology, Resistance to ARVs

Índice geral

AGRADECIMENTOS	II
RESUMO.....	III
ABSTRACT.....	IV
ÍNDICE DE FIGURAS	VIII
ÍNDICE DE TABELAS.....	X
LISTA DE ABREVIATURAS	XI
1. INTRODUÇÃO.....	1
1.1. O Vírus da Imunodeficiência Humana (VIH) e a Síndrome da Imunodeficiência Adquirida (SIDA)	1
1.2 Epidemiologia.....	1
1.3 Estrutura e Genoma do vírus	3
1.4 Ciclo de replicação do VIH-1	5
a) Ligação e Entrada	5
b) Transcrição reversa	6
c) Integração do ADN.....	7
d) Replicação viral	7
e) Maturação viral	9
1.5 Origem e diversidade genética.....	10
1.6 Distribuição geográfica dos subtipos, CRFs e URFs do VIH-1	14
1.7 Terapia Antirretroviral (TARV)	16

1.8	Mecanismos de Resistência aos ARVs	22
1.9	Caracterização sociodemográfica de São Tomé e Príncipe (STP)	25
1.10	VIH em São Tomé e Príncipe	26
2.	OBJETIVOS	28
2.1	Geral.....	28
2.2	Específicos	28
3.	MATERIAL E MÉTODOS.....	29
3.1	Casuística	29
3.2	Determinação dos subtipos do VIH-1	29
3.3	Análise Filogenética	30
3.3.1	Alinhamento e construção da árvore (inferência de árvore filogenética)	30
3.3.2	Filogeografia (reconstrução evolutiva e demográfica)	31
3.4	Avaliação da resistência aos ARVS (stanfordDB 8.7)	33
4.	RESULTADOS	35
4.1	Genotipagem (subtipagem).....	35
4.2	Análise Filogenética e Filogeográfica das sequências.....	36
4.3	Análise das Resistências aos ARVs.....	40
5.	DISCUSSÃO.....	47
5.1	Genotipagem.....	47
5.2	Filogenia e Filogeografia	50
5.3	Resistência aos ARVs.....	52
6.	CONCLUSÃO.....	60

7. REFERÊNCIAS BIBLIOGRÁFICAS	64
-------------------------------------	----

Índice de Figuras

Figura 1: Prevalência global de adultos dos 15 aos 49 infectados com o VIH em 2017	2
Figura 2: Estrutura do virião do VIH.....	4
Figura 3: Diagrama esquemático da organização dos genomas virais do VIH-1, VIH-2, SIV, stHIV-1 e SHIV.....	5
Figura 4: Representação esquemática do ciclo de replicação do VIH-1	9
Figura 5: Estrutura do VIH (a. Fase imatura b. Fase madura).....	10
Figura 6: Mapa de distribuição regional dos subtipos, CRFs e URFs do VIH-1, 2010-15 ..	15
Figura 7: Distribuição por região do número de pessoas em TARV, 2018.....	17
Figura 8: Visão geral do ciclo de replicação do VIH-1 e os mecanismos de ação dos ARVs	21
Figura 9: Mapa de São Tomé e Príncipe.....	26
Figura 10: Distribuição dos subtipos do VIH-1, CRFs e outros recombinantes circulantes em S.Tomé e Príncipe	35
Figura 11: Escala temporal da árvore de credibilidade máxima clade bayesiana (MCC) das estirpes VIH-1 CRF02_AG são-tomeneses e sequências globais próximas, extraídas do NCBI (n=187).....	37
Figura 12: Mapa ilustrativo da entrada da estirpe VIH-1 CRF02_AG em São Tomé e Príncipe.....	38
Figura 13: Representação do cluster de sequências Portuguesas e sequências de São-Tomé	39
Figura 14: Prevalência de resistências aos ARVs em estirpes de VIH-1 isoladas em S.Tomé e Príncipe	40
Figura 15: Distribuição dos padrões de resistência nas três classes de fármacos ARVs.....	42

Figura 16: Frequência das mutações de resistência consoante as estirpes do VIH-1 identificadas.....	43
Figura 17: Distribuição das percentagens dos níveis de resistência consoante as classes de fármacos ARVs.	44
Figura 18: Percentagem dos níveis de resistência relativamente aos antirretrovirais usados no tratamento do VIH-1 em S.Tomé e Príncipe.	45

Índice de Tabelas

Tabela 1: Resumo dos fármacos antirretrovirais.....	21
Tabela 2: Tabela de distribuição das frequências relativas e absolutas das mutações de resistência consoante as classes de ARVs (N=88).....	41
Tabela 3: Resumo das sequências, subtipos e mutações de resistência identificadas.....	46

Lista de abreviaturas

- 3TC:** Lamivudina
- ABC:** Abacavir
- ADN:** Ácido desoxirribonucleico
- ADNc:** ADN complementar
- aLTR:** Teste de razão de verossimilhança aproximada (do inglês: *likelihood ratio test*)
- ARN:** Ácido ribonucleico
- ARNm:** ARN mensageiro
- AZT:** Zidovudina
- BEAST:** Programa Bayesiano de Análise Evolutiva por Árvores de Amostragem
- BEAUti:** Programa Utilitário de Análise Evolucionária Bayesiana
- BP:** Par de bases (do inglês: *base pair*)
- CA:** Cápside viral (p24)
- CCR5:** Receptor de quimiocina 5
- CD4:** *Cluster de diferenciação 4* (do inglês: *cluster of differentiation 4*)
- CDC:** Centros de Controlo e Prevenção de Doenças (do inglês: *Centers for Disease Control and Prevention*)
- CPLP:** Comunidade dos países de língua oficial portuguesa
- CRF:** Formas Recombinantes Circulantes (do inglês: *circulating recombinant form*)
- CRXC4:** Receptor de quimiocina CRXC4
- d4T:** Estavudina
- DRM:** Multirresistência a medicamentos (do inglês *Drug Multiresistance*)

ENV: Gene que codifica as proteínas do invólucro viral

ESS- Tamanhos de amostra eficazes (do inglês: *Effective Sample Sizes*)

FDA: Federação de Administração de Alimentos e Medicamentos (do inglês: *Food and Drug Administration*)

FTC: Emtribicina

GAG: Antígenos específicos do grupo (do inglês: *group specific antigens*)

GRT: Teste de Resistência Genotípicas (do inglês: *Genotypic Resistance Test*)

HAART: Terapia Antirretroviral Altamente Ativa (do inglês: *highly active antiretroviral therapy*)

HKY: (do inglês: *Hasegawa-Kishino-Yano*)

HIVDB: Base de dados do VIH (inglês: *HIV Databases*)

HIVDR: Resistência às drogas contra o HIV (do inglês: *HIV Drug Resistance*)

HR1: hepteto repetido 1 (do inglês: *heptad repeat 1*)

HR2: hepteto repetido 2 (do inglês: *heptad repeat 2*)

IDH: Índice de Desenvolvimento Humano

IF: Inibidor de Fusão

IN: Integrase

ISTs: Infecções Sexualmente Transmissíveis

LTR : Repetição de terminação longa (do inglês: *Long terminal repeat*)

MA: Matriz/membrana central externa (p17)

MCC: Clado de máxima credibilidade (do inglês: *Maximum credibility clade*)

MCMC: (Do inglês: *Markov Chain Monte Carlo*)

Mg: Miligramas

ML: Máxima Verossimilhança (do inglês: *Maximum Likelihood*)

MVC: Maraviroc

NAM: Mutações associadas a nucleósidos / nucleosídeos

NRTIs: Inibidores da transcriptase reversa nucleósidos

NNRTIs Inibidores da transcriptase reversa não nucleósidos

IN/ InSTIs: Inibidores de integrase/ inibidores da transferência da cadeia de integrase

ODS: Objetivos do Desenvolvimento Sustentável

OMS: Organização Mundial da Saúde

NC: Nucleocapsídeo viral (p7)

Nef: Fator regulador negativo

Nm: Nanomêtro

PEP/PreP: Profilaxia pós-exposição

PI: Inibidores de Protease (do inglês: *Protease inhibitor*)

PIB: Produto Interno Bruto

PNLS: Programa Nacional de Luta contra a SIDA

Pol: Gene da polimerase viral

PNUD: Programa das Nações Unidas para o Desenvolvimento

PR: Protease viral

PreP: Profilaxia Pré-Exposição

PP: Probabilidade posterior (do inglês: *posterior probability*)

PSP: Probabilidade de estado posterior (do inglês: *posterior state probability*)

PVHS: Pessoas que vivem com o VIH/SIDA

REV: Regulador de junção (do inglês: *splicing-regulator*)

RT: Transcriptase Reversa (do inglês: *Reverse Transcriptase*)

SDRM: Vigilância de mutações de resistência (do inglês: *Surveillance Drug Resistance Mutation*)

SIDA: Síndrome da Imunodeficiência Humana

SIV: Vírus da Imunodeficiência Símia (do inglês: *Simian immunodeficiency vírus*)

SIVsmm Vírus da imunodeficiência símia do sooty mangabey (do inglês: *Simian immunodeficiency virus from sooty mangabey monkeys*)

SIVcpz: Vírus da imunodeficiência símia dos chimpanzés (do inglês: *Simian immunodeficiency virus from chimpanzés*)

SPREAD: Reconstrução filogenética espacial da dinâmica evolucionária (do inglês: *Spatial Phylogenetic Reconstruction of Evolutionary Dynamics*)

STP: São Tomé e Príncipe

TAM: Mutações análogas da timidina (do inglês: *Thymidine analogue mutation*)

TAR: Sequência alvo para transativação viral (do inglês: *Target sequence for viral transactivation*)

TDF: Tenofovir

TDR: Resistência transmitida a fármacos (do inglês: *Transmitted Drug Resistance*)

UNAIDS: Programa Conjunto das Nações Unidas sobre VIH / SIDA (do inglês: *Joint United Nations Programme on HIV/AIDS*)

URF: Formas Recombinantes Únicas (do inglês: *Unique Recombinant Form*),

VIH: Vírus da Imunodeficiência Humana

Vif: Fator de infetividade viral

VL: Carga viral (do inglês: *Viral Load*)

Vpr: Proteína viral R (do inglês: *Viral Protein R*)

Vpu: Proteína viral U (única) (do inglês: *Viral Protein U*)

Vpx: Proteína do virião encontrada no HIV-2

1. Introdução

1.1. O Vírus da Imunodeficiência Humana (VIH) e a Síndrome da Imunodeficiência Adquirida (SIDA)

O VIH é o agente responsável pela Síndrome da Imunodeficiência Adquirida (SIDA), definida como uma forma grave de imunodeficiência humana, que geralmente envolve uma redução no número de linfócitos T auxiliares (CD4), e é acompanhada por múltiplas infecções oportunistas ou malignidades (1).

O VIH é um vírus pertencente a família dos Retrovírus (*Retroviridae*), subfamília *Orthoretrovirinae* e agrupado ao gênero *Lentivirus* (2). Este está dividido em dois tipos, o VIH-1 e o VIH-2, baseado nas suas características genéticas e diferenças nos antígenos virais (2,3).

A SIDA foi inicialmente identificada em 1981 pela *Centers for Disease Control and Prevention* (CDC) em homens gays e utilizadores de drogas endovenosas (4). Apenas em 1983 foi isolado o agente causal desta síndrome, o VIH-1. A infecção pelo vírus caracteriza-se por ser crónica e incurável, porém pode ser tratada com combinações de medicamentos, designados por terapia antirretroviral altamente ativa (HAART) (5).

1.2 Epidemiologia

A transmissão VIH continua a ser um grande problema de saúde global, pois já causou a morte de mais de 35 milhões de pessoas desde a sua descoberta até os dias atuais (6).

Em 2017, dados da OMS (Organização Mundial da Saúde) apontaram para cerca de 940 mil mortes em todo o mundo de causas relacionadas com o vírus, além de aproximadamente 36,9 milhões de pessoas a viver com o VIH até o final de 2017, e cerca de 1,8 milhões de pessoas infetadas em todo o mundo no mesmo período (6,7).

A Região Africana da OMS é a mais afetada pela epidemia, com cerca de 25,7 milhões de pessoas a viver com o vírus em 2017, representando mais de dois terços do total global de novas infecções (Figura 1) (6,8).

A maior carga de doença e o maior risco de novas infecções localiza-se nas populações-chave: grupos de pessoas que estão em risco aumentado de ter VIH, independentemente do tipo de epidemia ou do contexto local (6). Estas populações apresentam frequentemente questões legais e sociais relacionadas com os seus comportamentos, além de vulnerabilidade aumentada ao vírus, e acesso reduzido a programas de teste e tratamento (6). Tais populações incluem trabalhadores do sexo, homens que têm sexo com homens, pessoas que injetam drogas, transsexuais e pessoas em prisões e/ou outros ambientes fechados (6,7). Estima-se então que 47% de novas infecções ocorram neste grupo e em seus parceiros sexuais (6,7).

As mulheres estão entre os grupos mais afetados pelo vírus em algumas regiões do globo, comparativamente aos homens (7). Este hiato de gênero é particularmente notável na África subsaariana onde 56% das pessoas a viver com o VIH são mulheres (7). Contudo, os índices de redução da mortalidade nas pessoas que vivem com o VIH são maiores nas mulheres comparativamente com os homens, refletindo uma maior cobertura de tratamento entre as mulheres (7).

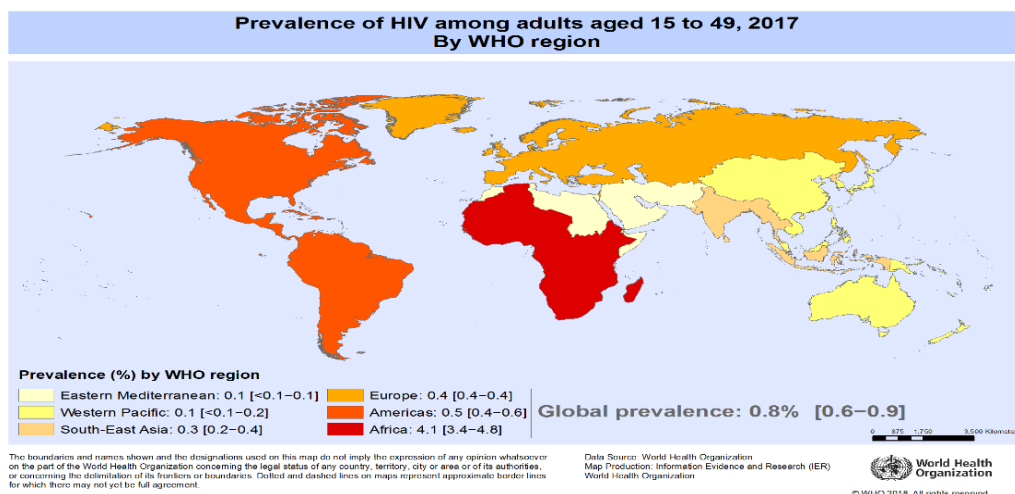


Figura 1: Prevalência global de adultos dos 15 aos 49 infectados com o VIH em 2017

Fonte: <https://www.who.int/gho/hiv/en/>

1.3 Estrutura e Genoma do vírus

O genoma dos retrovírus é único entre os vírus em vários aspectos, incluindo sua organização, seu modo de síntese e suas funções na replicação (9).

A partícula viral do VIH é esférica e tem um diâmetro de aproximadamente 120 nm. Está envolvida por uma membrana de bicamada lipídica externa, rica em lipoproteínas, que constitui o seu envelope (10). Esta membrana lipídica (envelope), contém complexos de trómeros de superfície, as glicoproteínas gp120 (proteína de superfície (SU)) ligadas aos trómeros transmembranares (proteínas transmembranares (TM)), gp41, que estão ancoradas a estrutura do envelope (Figura 2) e (Figura 5) (11).

O genoma do VIH é composto por duas cópias de moléculas de ARN positivo de cadeia simples (Figura 2) (2,11). O material genético está envolvido pela cápside viral que é composta pela proteína p24, que por sua vez encontra-se ancorada pelo polo cónico a proteína da matriz (p17) (Figura 2) e (Figura 3) (2,10).

O genoma do VIH contém nove genes que codificam quinze proteínas virais (Figura 3) (12). Três genes principais, *gag*, *pol* e *env*, codificam proteínas estruturais (matriz, cápside, nucleocápside e p6), enzimas virais (protease, transcriptase reversa [RT] e integrase) e proteínas de envelope (gp120 e gp41), respetivamente (Figura 2) e (Figura 3) (2,12,13). Os genes restantes codificam para proteínas reguladoras (Tat e Rev) e proteínas acessórias (*Vif*, *Vpu* / *Vpx*, *Vpr* e *Nef*) (Figura 3) (3). O *vpu* é encontrado exclusivamente no VIH tipo 1 (VIH-1), enquanto o *vpx* é encontra-se no VIH-2 (Figura 3) (12,14).

As proteínas reguladoras Tat (proteína transativadora) e Rev (*RNA-splicing-regulator*) são necessárias para o início da replicação do vírus, enquanto as outras proteínas reguladoras Nef (fator regulador negativo), Vif (fator de infectividade viral), Vpr (proteína viral r) e Vpu (proteína viral única) têm um impacto na replicação viral, na maturação do vírus e na patogénese, e o Vpx é parcialmente responsável pela reduzida patogenicidade do VIH-2 (Figura 2) e (Figura 3) (14).

Index to 3D Views of HIV Macromolecular Structures

Clickable LINKS are underlined

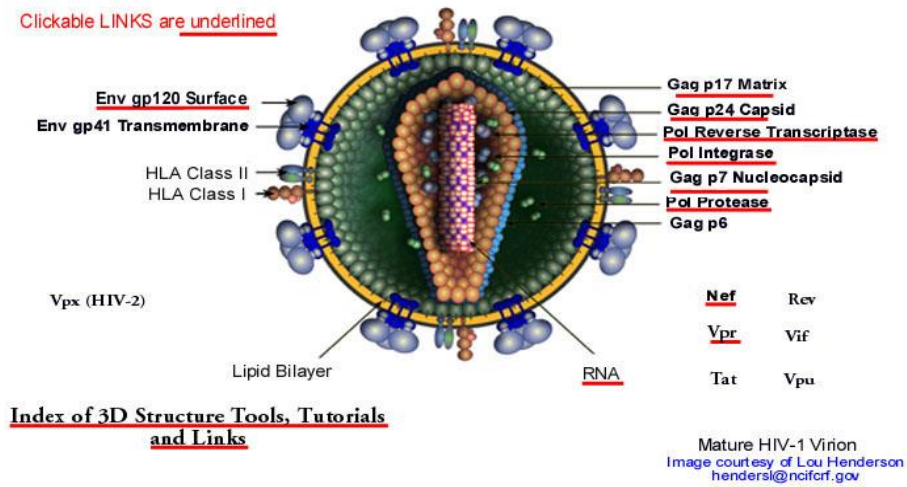


Figura 2: Estrutura do viri o do VIH.

Fonte: <https://www.hiv.lanl.gov/content/sequence/STRUCTURE/INDEX.HTML>

Figure 2

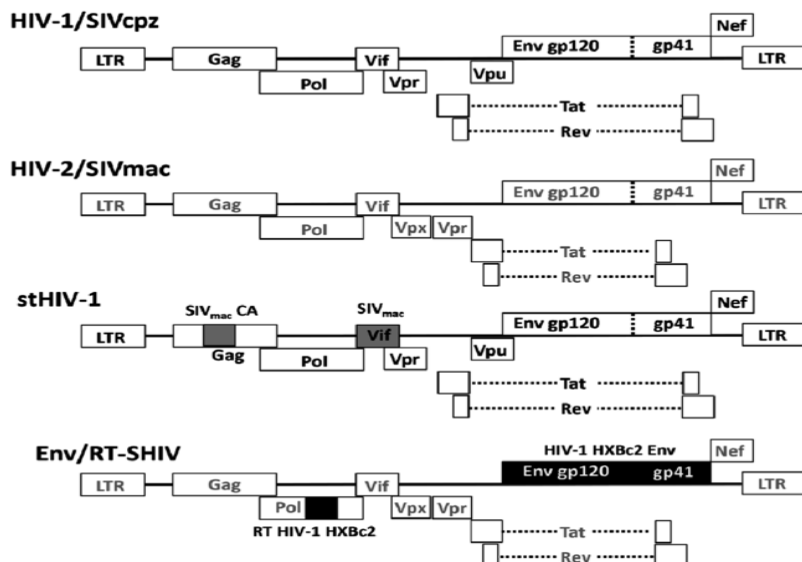


Figura 3: Diagrama esquemático da organização dos genomas virais do VIH-1, VIH-2, SIV, stHIV-1 e SHIV. Abreviações: SIV, vírus da imunodeficiência símia; stHIV-1, vírus da imunodeficiência humana símio-tropical; SHIV, vírus da imunodeficiência símio-humano; SIVcpz, derivado de chimpanzé; SIVmac, derivado de macaque. Fonte: artigo: Nonhuman Primates and Humanized Mice for Studies of HIV-1 Integrase Inhibitors: A Review: Said A. Hassounah et al.

1.4 Ciclo de replicação do VIH-1

O ciclo de replicação do VIH pode ser resumido em cinco etapas; 1) ligação e entrada; 2) transcrição reversa; 3) integração do ADN; 4) replicação viral; 5) maturação viral (Figura 4).

a) Ligação e Entrada

A entrada do vírus pode ser dividida em três grandes eventos: ligação do vírus à célula, ativação e fusão (10).

Os passos iniciais dessa infecção caracterizam-se por interações complexas proteína-proteína, sendo o complexo de trómeros do envelope viral (gp160), essencial para o reconhecimento e entrada do vírus na célula-alvo (2,10).

Células hospedeiras como as células T auxiliares, monócitos, macrófagos, células dendríticas, eosinófilos, células da micróglia do sistema nervoso central (astrócitos), que expressam na superfície celular a glicoproteína CD4 (*cluster of differentiation 4*), são os alvos principais dos “botões” do envelope do VIH (2,10,14).

O início da infecção dá-se pela interação da gp120 com a molécula CD4, e um recetor de quimiocina geralmente o CCR5 ou CXCR4 na superfície da célula (Figura 4) (15) .

Há uma alteração tridimensional da molécula gp120 ao ligar-se ao co-receptor, o CCR5 (receptor de quimiocina 5) ou o CXCR4 (receptor de quimiocina CRXC4) exhibe o peptídeo de fusão (gp41), possibilitando a fusão do envelope viral a membrana da célula hospedeira (Figura 4) (2,10). Após a ligação a membrana celular as moléculas gp120 e gp41 dissociam-se (16), e o peptídeo de fusão N-terminal hidrofóbico da gp41 inicia a fusão do vírus na membrana celular (Figura 4) (2,15).

Em seguida, as sequências de repetição na gp41, HR1(*heptad repeat 1*) e HR2 (*heptad repeat 2*) interagem entre si, provocando o colapso da porção extracelular da gp41 num *hairpin*, estrutura circular que aproxima o vírus da membrana celular, permitindo a fusão das suas membranas e consequente entrada da cápside viral (2,10,15) (17).

Após a fusão, a cápside viral é transportada para o citoplasma, e induz a liberação do conteúdo da cápside no citoplasma da célula alvo, do ARN viral (não revestido) e das enzimas virais (Figura 4) (2,10).

b) Transcrição reversa

A conversão do ARN viral em ADN viral ocorre devido a ação da transcriptase reversa, e a ativação desta enzima ocorre no citoplasma (2). Este processo de transcrição é predisposto a erros, ocasionando mutações que podem causar resistência aos antirretrovirais (2).

O genoma do ADN viral, é gerado pela transcrição reversa do ARN viral em ADN, degradação do ARN e integração do ADN viral de cadeia dupla no genoma humano (Figura 4) (2). Para que isto aconteça, a TR liga-se ao local iniciador do ARN viral na extremidade

5', no citoplasma, e inicia a transcrição reversa e polimerização da fita negativa, transformando-a em ADN de cadeia simples ou cADN (ADN complementar), até que o ARN viral seja transcrito numa hélice híbrida de ARN / ADN (Figura 4) e (Figura 8) (2,10).

Paralelamente à síntese do cADN, a cadeia de ARN é degradada enzimaticamente pela atividade RNase H da RT. De seguida ocorre a conversão do cADN de cadeia simples em ADN de cadeia dupla (ADN pró-viral) pela atividade da ADN-polimerase dependente da ADN da RT (2).

c) Integração do ADN

A RT fragmenta os nucleótidos de cada extremidade 3' do ADN de dupla hélice, criando duas extremidades aderentes (ADN do pró-vírus) que é transportado para o núcleo da célula na forma de um complexo constituído pela integrase (IN) e ADN pró-viral linear (10). A integrase então insere aleatoriamente o genoma do pró-vírus no genoma da célula hospedeira humana, e ali finaliza a infeção do VIH na célula hospedeira e o estabelecimento de uma infeção persistente (Figura 4) e (Figura 8) (2).

As enzimas do hospedeiro completam o processo de integração reparando as lacunas da cadeia simples adjacentes às extremidades do DNA viral unidas, resultando no estabelecimento de um pró-vírus estável (18). Contudo, a integração do ADN pró-viral e a expressão do pró-vírus requerem que a célula alvo esteja em um estado ativo (10). Monócitos e macrófagos, células da microglia e células T CD4 + quiescentes latentemente infetadas contêm pró-vírus integrados e são importantes reservatórios celulares de longevidade do VIH na célula (10).

d) Replicação viral

A biogênese, transporte e “tradução” do ARN mensageiro marca a transição da fase inicial para a tardia da replicação viral, na qual o foco muda para a expressão genética, seguida pela montagem e saída de partículas virais nascentes (18).

Para que ocorra a transcrição do ADN pró-viral para o ARN mensageiro (mARN) é necessário que a célula hospedeira esteja ativa, e além disso, são necessários fatores de

transcrição que permitem iniciar o processo de transcrição pela ARN polimerase (Figura 4) (19,20).

Após a transcrição do ADN para o ARN mensageiro (mARN), o mARN é reorganizado por meio de *splicing*, que consiste no processo de maturação do mARN, na qual regiões não codificantes (os intrões) são retirados do mARN, deixando somente as regiões codificantes (os exões) (10). De seguida este mARN é transportado do núcleo para o citoplasma, onde proteínas estruturais de novos viriões são sintetizadas (Figura 4) (10). Inicialmente ocorre a síntese de proteínas reguladoras do VIH-1, tais como Tat (são proteínas transativadoras transcricionais do LTR: incentiva a produção de novos vírus) que liga-se ao elemento TAR (Elemento de Resposta à Transativação) no início do ARN do VIH-1 e a proteína Rev (está envolvida no encerramento dos ARNs do núcleo e da célula) (10,20).

A expressão genética é sequencial, e é controlada por Tat e Rev, sendo que a maioria dos mRNAs virais encontrados no citoplasma já sofreu vários eventos de *splicing* e codifica principalmente para Tat, Nef e Rev (20).

A região 5' LTR contém o promotor para a transcrição dos genes virais (2). Na direção 5' a 3' segue-se a grelha de leitura e tradução dos genes virais, sendo que o *gag*, codifica para as proteínas da membrana central externa (MA, p17), a proteína da cápside (CA, p24), a nucleocápsideo (NC, p7), seguido pelo gene *pol* que codifica para as enzimas protease (PR, p12), transcriptase reversa (RT, p51) e ribonuclease H (p15) ou RT mais ribonuclease H junto (p66) e integrase (IN, p32) (Figura 3) (2). Adjacente ao gene *pol*, segue-se a estrutura de leitura do *vif/vpr*, e depois o gene *env* que codifica para as glicoproteínas do envelope, gp120 (proteína de superfície, SU) e gp41 (proteína transmembranar, TM)) (Figura 3) (2,10).

A fragmentação das moléculas precursoras (proteína Env, poliproteína Gag (p55) e Pol) pela protease do VIH-1 é necessária para a geração de partículas virais infecciosas (Figura 4) (10).

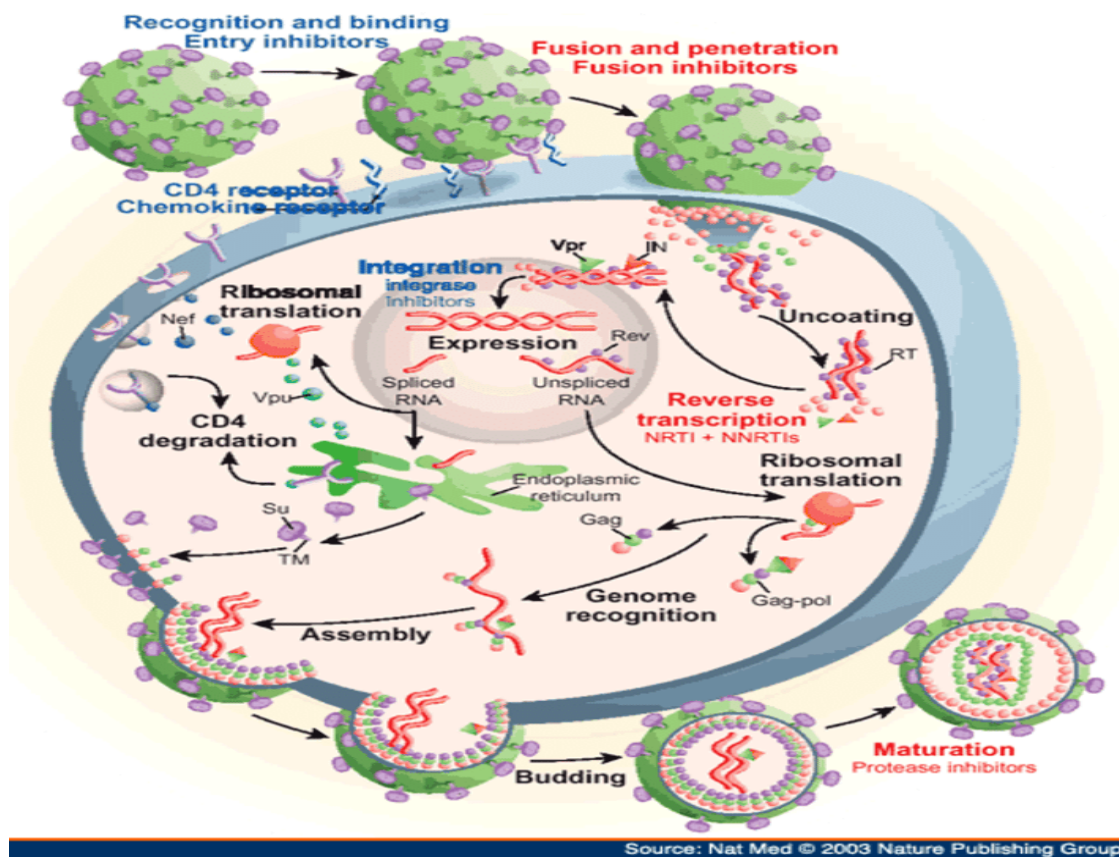


Figura 4: Representação esquemática do ciclo de replicação do VIH-1. Fonte: Pomerantz e Horn, 2003; Turner e Summers, 1999.

e) Maturação viral

A maturação da partícula do VIH ocorre pela fragmentação das proteínas precursoras Gag e Gag / Pol em proteínas individuais pela protease (2), para produzir os componentes estruturais MA, CA e NC e p6, e as enzimas PR, RT e IN (18), que resultam em novas partículas virais infecciosas, e que se gemulam através da membrana da célula hospedeira, adquirindo assim um novo envelope (Figura 4) (10,15).

Durante a morfogénese do vírus na célula infetada, este pode também incorporar na sua membrana diferentes proteínas da membrana da célula hospedeira, como as proteínas HLA

de classe I e II, ou proteínas de adesão como a ICAM-1, e isso pode facilitar a adesão a outras células-alvo (Figura 5) (10).

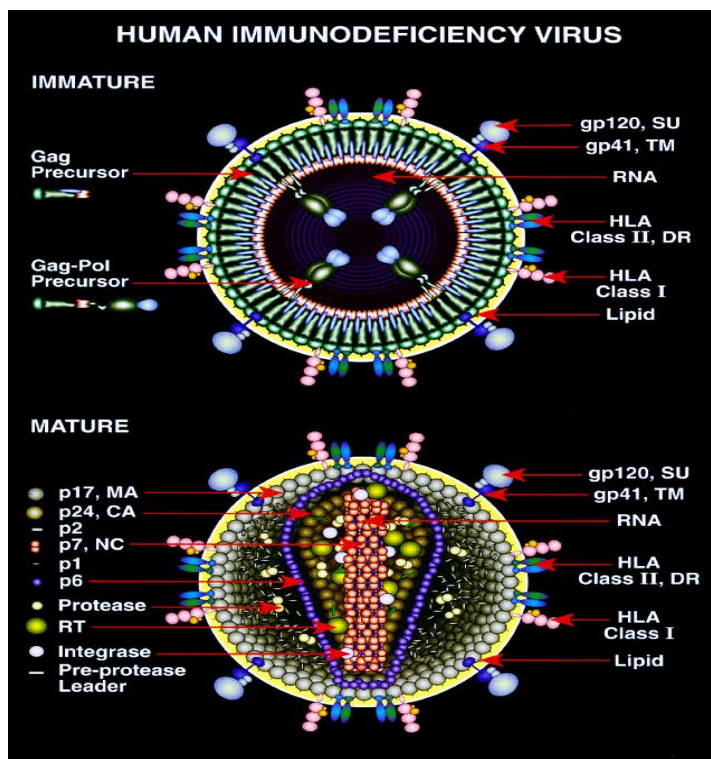


Figura 5: Estrutura do VIH (a. Fase imatura b. Fase madura)

Fonte: <https://www.aidsreagent.org/program>

1.5 Origem e diversidade genética

O VIH ter-se-á originado de múltiplas transmissões zoonóticas do vírus da imunodeficiência simia (SIV) de primatas não humanos para humanos na África Ocidental e Central (9,21).

Evidências filogenéticas demonstraram que a origem do vírus está relacionada com primatas africanos, tendo o VIH-1 sido transmitido a partir do SIV de chimpanzés da subespécie *Pan troglodytes troglodytes* SIV (SIVcpz), e o VIH-2 a partir do SIV de macacos *sooty mangabey* (SIVsmm) (11). A análise por sequenciamento genético revelou que os genomas do SIVsmm

(*sooty mangabey SIV*) e do VIH-2 exibem um alto grau de similaridade, e o SIVcpz (vírus da imunodeficiência dos chimpanzés) é mais próximo do VIH-1) (Figura 3) (2,9). Assume-se por isso que estes dois vírus resultaram de eventos distintos de transmissão entre espécies (22).

Diversas análises epidemiológicas e filogenéticas indicam que o VIH-1 foi transmitido para a população humana no final do século XIX/início do século XX, tendo a epidemia começado o seu crescimento exponencial na República Democrática do Congo na primeira metade do século XX, coincidentemente com várias alterações industriais, comportamentais e sociais (23,24).

Nos EUA, em 1981, foram descritos os primeiros casos da síndrome da imunodeficiência adquirida (SIDA) em homens homossexuais (4,11), tendo sido confirmado posteriormente por cientistas franceses (Luc Montagnier e Barre Sinoussi) (25) e americanos (Robert Gallo) (1) em 1983 e 1984, que o Retrovírus VIH-1 era o agente causador da epidemia.

O VIH apresenta uma notável capacidade de mutação e adaptação às novas condições do ambiente humano, e possui uma extraordinária diversidade genética global (1,9,10). Este facto deve-se a uma elevadíssima taxa de replicação e à falta de uma atividade de revisão da enzima RT viral (1), que leva a uma grande incidência de erros a nível da transcrição e mutações nos nucleótidos que compõem o genoma viral, durante o seu ciclo de replicação (1,9). Além disso, vários mecanismos de erro atribuídos às polimerases, geram uma alta frequência de mutações como substituições genéticas, deleções, recombinações, repetições e inserções, podendo envolver um ou mais nucleótidos (9).

Com base nas características genéticas, nas diferenças nos antígenos virais, nas relações filogenéticas, características clínicas, virulência, e infecciosidade, o VIH é classificado nos tipos 1 e 2 (VIH-1, VIH-2) (2,11).

Embora a estrutura básica (genes estruturais: *gag*, *pol*, *env*, *tat*, *rev*, *vif*, *nef* e *vpr*) seja a mesma para todos os retrovírus, o VIH-1 e VIH-2 apresentam diferenças na organização do seu genoma, e uma semelhança de nucleótidos de apenas 60%, (7,10). O agente principal da

SIDA a nível mundial é o VIH-1, comparativamente com o VIH-2 cuja ocorrência está restrita a algumas regiões da África Ocidental, Central e Europa ocidental (9,26).

O VIH-1 apresenta estirpes circulantes que possuem um grau extraordinário de diversidade genética, e que podem influenciar aspetos da sua biologia, tais como infecciosidade, transmissibilidade e imunogenicidade (27).

As sequências derivadas das estirpes do VIH-1 foram historicamente classificadas com base nas suas relações filogenéticas em grupos e subtipos (27).

As complexidades crescentes das sequências do VIH-1 indicaram a necessidade de revisão do sistema de nomenclatura, e assim em setembro de 1999, no Instituto Santa Fé, Novo México, foram definidas as particularidades da nomenclatura do VIH-1 agora em uso (27).

A classificação atual do VIH-1 está feita com base na divisão deste em quatro grandes grupos ou subtipos: M (main ou principal), O (outlier), e o N (não-M / não -O) e P (2,9,27).

Na linhagem dos lentivírus, os grupos M e N e SIVcpz são aproximadamente equidistantes entre si, enquanto o grupo O é o mais distante das outras estirpes (9).

Após a transmissão zoonótica de primatas não humanos para humanos no início do século 20, o grupo M do VIH-1 diversificou-se em subtipos distintos, designados pelas letras A, B, C, D, F, G, H, J, K e L (28), além dos CRFs (Formas Recombinantes Circulantes) que correspondem a recombinantes que circulam em várias pessoas não epidemiologicamente relacionadas e que circulam em diferentes extensões em populações ao redor do globo. Existem ainda Formas Recombinantes Únicas (URFs), que correspondem a formas circulantes que circulam apenas num grupo restrito de pessoas epidemiologicamente relacionadas entre si (9,29).

Embora apenas os vírus do grupo M tenham estabelecido disseminação pandêmica, outras transmissões separadas do SIV para seres humanos na bacia do rio Congo levaram à transmissão não-pandêmica dos grupos VIH-1 O, N e P (23).

Os CRFs (Formas Recombinantes Circulantes) referem-se às estirpes do VIH-1 que exibem ancestrais comuns discordantes nas filogenias, inferidas em diferentes partes do seu genoma, sugerindo fortemente que estes vírus sejam produto de eventos de recombinação de dois ou

mais subtipos (27). Todos os CRFs da mesma linhagem devem compartilhar uma estrutura de mosaico idêntica, confirmando que são descendentes do mesmo evento de recombinação (2), (27).

Estas estirpes estão propagadas na população e são nomeados consecutivamente de acordo com as diretrizes acordadas internacionalmente, descrevendo uma linhagem recombinante que desempenha um papel importante na pandemia do VIH-1 no mundo (29) (27). Até agora foram identificados 97 CRFs pertencentes ao grupo M do VIH-1 e 1 CRF pertencente ao VIH-2 (30), sendo que os CRFs do VIH-1 mais importantes são os CRF01_AE e CRF02_AG.

As recombinações em que os vírus têm origens filogenéticas discordantes em diferentes regiões genômicas, mas não se assemelham a quaisquer CRFs previamente descritos, e sem evidências de transmissão posterior, são definidas como Formas Recombinantes Únicas (URFs) (27,29). Existem URFs envolvendo mais de dois subtipos ou também incluindo fragmentos genômicos do CRF (9). A maioria dos URFs é detetada em regiões onde múltiplos subtipos e CRFs co-circulam (9).

Para a definição de um novo subtipo, sub-subtipo ou CRF, as estirpes devem ser identificadas em pelo menos três indivíduos sem ligação epidemiológica direta que partilham a mesma estrutura recombinante e formam um agregado monofilético em todas as regiões do genoma (9,27). Para isso, são usadas três sequências genômicas quase completas, ou dois genomas completos em conjunto com sequências parciais de uma terceira estirpe e devem se assemelhar entre si, mas não a outros subtipos existentes em todo o genoma (9,27).

Grupos monofiléticos são definidos como grupos contendo espécies que estão mais intimamente relacionadas entre si do que com qualquer outro grupo da árvore filogenética (9).

Genomas inteiros ou suas regiões parciais que formam linhagens distintas em relação aos subtipos conhecidos, e não atendem aos critérios para designar um novo subtipo ou sub-subtipo, serão rotulados como não classificados (U: *Unclassified*) (27). Segmentos dentro de genomas recombinantes inter-subtipo, incluindo segmentos dentro de CRFs, para os quais a

estirpe "parental" não pode ser determinada, serão marcados com "U", isto é, um subtipo só pode ser definido para uma linhagem inferida como não recombinante (27).

1.6 Distribuição geográfica dos subtipos, CRFs e URFs do VIH-1

O padrão de distribuição mundial do VIH-1 sofreu mudanças que têm vindo a ser estudadas ao longo do tempo, tendo vindo a ser demonstrado que as proporções dos subtipos do grupo M do VIH-1, os CRFs e os URFs diferiram entre as regiões ao longo do tempo. A diversidade maior localiza-se na África Central, onde foram encontrados todos os subtipos de VIH-1 e muitos dos CRFs e URFs, em todos os períodos de tempo (29).

Com o tempo, diminuições nas proporções dos subtipos A, D, G e H foram acompanhadas por aumentos nas proporções do subtipo C, outros CRFs e URFs (29).

Dentre os subtipos do grupo M do VIH-1, o subtipo C dominou e continua a dominar a epidemia mundial (Figura 6), embora este perfil tenha se alterado um pouco. Este subtipo é o mais prevalente globalmente, e representa 46,6% das PVHS (pessoas que vivem com o VIH_SIDA) no período de 2010-15 (29), ocorrendo principalmente nos países da África Austral e Oriental e sul da Ásia (Índia), onde contribuiu com pelo menos 89% das infeções ao longo do período de 1990-2015 (Figura 6) (2,29).

Já o subtipo B foi responsável por 12,1% das infeções mundiais, estando amplamente disperso globalmente, sendo a linhagem dominante na Europa Ocidental e Central, nas Américas (América do Norte, Caribe, América Latina) e Oceânia, onde foi responsável por pelo menos 75% das infeções de 2010 a 2015 (Figura 6) (2,29).

O terceiro subtipo com importância epidemiológica e distribuição global é o subtipo A, com 10,3% das infeções por VIH-1, com predomínio na Europa Oriental, Ásia Central e África Oriental, com mais de 50% das infeções causadas em todos os períodos do tempo (Figura 6) (9,29).

Além destes subtipos, os CRFs também têm grande contribuição nas infeções por VIH-1, sendo que o CRF02_AG representa 7,7% das infeções, ocorrendo em grande proporção na

África Ocidental 46,2%, no período de 2010-2015, seguido por CRF01_AE 5,3%, que predomina no sudeste e leste da Ásia, com uma proporção global de recombinantes de cerca de 80% em cada região de 2010 a 2015 (Figura 6) (29).

Os outros CRFs representam 3,7% das infecções, e com isto eleva-se assim a proporção total das infecções por CRFs para 16,7% (29). Os URFs representaram 6,1% das infecções, o que elevou a proporção de todas as infecções globais por VIH-1 atribuíveis aos recombinantes a 22,8% (Figura 6) (29).

O subtipo G representa 4,6% das infecções e o subtipo D 2,7% (29). Além disso, os subtipos F, H, J e K combinados representaram 0,9% das infecções mundiais (Figura 6) (29).

As infecções do Grupo O são menos comuns, e têm sido endêmicas no oeste da África Central, principalmente nos Camarões, Gabão, Nigéria e Guiné Equatorial, tendo sido responsável por menos de 10% das infecções por HIV-1 em todo o mundo (9). Este grupo já foi identificado na África e espalhou-se na Europa e nos Estados Unidos, no entanto, a maioria das infecções por grupo O pode estar diretamente ligada a pessoas que tiveram alguma conexão com a África Central Ocidental (9).

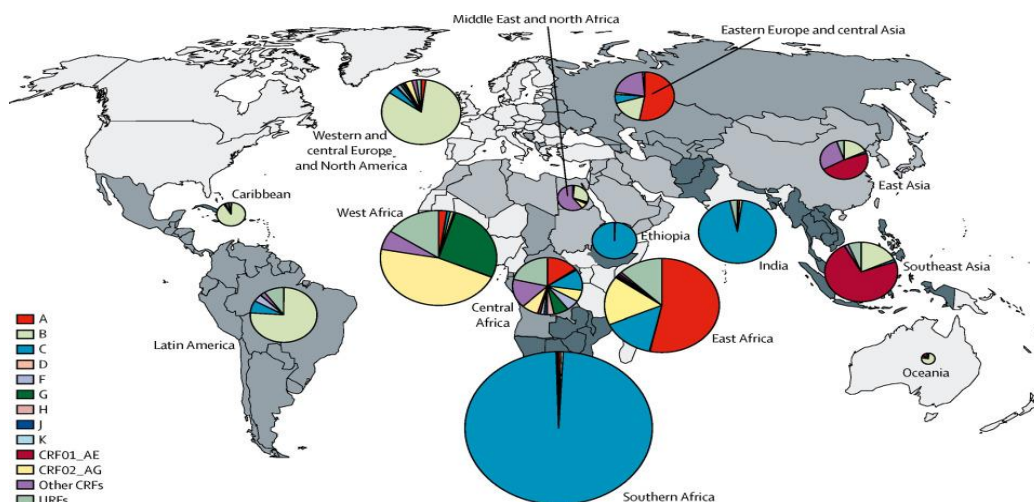


Figura 6: Mapa de distribuição regional dos subtipos, CRFs e URFs do VIH-1, 2010-15

Fonte: Global and regional molecular epidemiology of HIV-1, 1990–2015: a systematic review, global survey, and trend analysis; Joris Hemelaar et al. 2019

1.7 Terapia Antirretroviral (TARV)

Não há cura para a infecção pelo VIH, no entanto, medicamentos antirretrovirais (ARV) eficazes podem ajudar a controlar a replicação do vírus e a prevenir a sua transmissão, permitindo que pessoas portadoras, e aquelas em risco substancial de adquirir a infecção, possam desfrutar de uma vida longa, saudável e produtiva (6).

O número anual de mortes por doenças relacionadas à SIDA entre pessoas que vivem com o VIH, caiu de um pico de 1,9 milhões em 2004 para 940 mil em 2017 (7). Sendo estas reduções em grande parte impulsionadas pelo maior acesso à terapia antirretroviral, mas não com a rapidez suficiente para alcançar o marco da Assembleia Geral de 2020 (7). Em 2017, 21,7 milhões de pessoas a viver com o VIH estavam a receber terapia antirretroviral (TARV) globalmente (Figura 7) (6).

O número de novas infeções pelo vírus continuou a diminuir globalmente em 2017, e estimativas modeladas mostram que as novas infeções, para todas as idades, diminuiram de um pico de 3,4 milhões em 1996 para 1,8 milhões em 2017 (7).

A TARV consiste numa combinação de três ou mais medicamentos direcionados a diferentes fases do ciclo de vida do VIH (Figura 8) com o objetivo de suprimir a replicação viral e reduzir a carga viral plasmática deste abaixo dos limites de deteção dos ensaios clínicos mais sensíveis (50 cópias do RNA viral / ml), e preservar ou restaurar a função imunológica (31, 33).

As principais combinações de fármacos da terapia antirretroviral altamente ativa (HAART), são direcionados contra a atividade enzimática da RT, PR e da IN, e suprimem as cargas virais e as taxas de transmissão, entretanto podem gerar complicações relacionadas com a toxicidade destes compostos e a selecção de estirpes resistentes (18).

Com o início da TARV e a ampliação da sua cobertura para vários países do mundo, o número de novas infeções pelo VIH caiu em cerca de 36%, e as mortes relacionadas com o mesmo reduziram em 38% no período de 2000 a 2017 (6). Com isto cerca de 11,4 milhões de vidas foram salvas devido à TARV neste período (6).

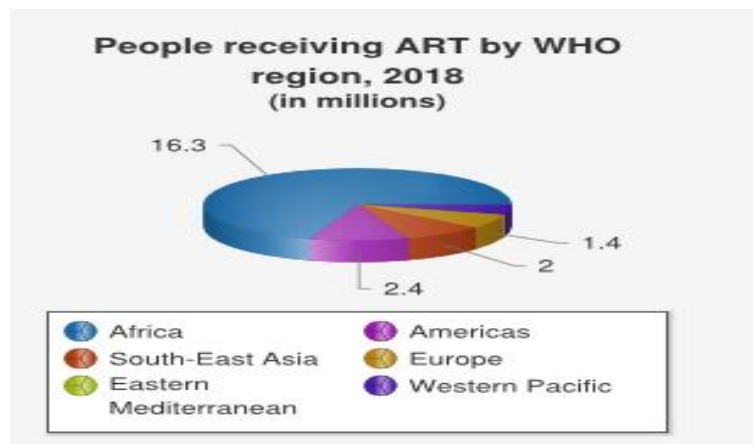


Figura 7: Distribuição por região do número de pessoas em TARV, 2018.

Fonte: World Health Organization 2018: <http://apps.who.int/iris>

As diretrizes consolidadas da OMS de 2016 sobre o uso de medicamentos antirretrovirais para o tratamento e prevenção da infecção pelo VIH são baseados em princípios de disponibilidade, acessibilidade, aceitabilidade e qualidade, cuja abordagem é centrada na atenção às pessoas, focada e organizada em torno das necessidades, preferências e expectativas de saúde de pessoas e comunidades, sustentando a dignidade individual, a equidade, o respeito e a autonomia (34). E com isso, recomendam testar e tratar todas às pessoas com diagnóstico de VIH, independentemente da contagem de CD4, visando melhorar a qualidade no tratamento do VIH e aproximar-nos dos ideais universais de serviços integrados de cobertura de saúde, abordagens de cuidados de saúde centradas na comunidade, e responsabilidade compartilhada (31).

Expandir o acesso ao tratamento está no centro das novas metas de tratamento para 2020 com o objetivo de acabar com a epidemia da SIDA como uma ameaça à saúde pública até 2030 (31). O objetivo 90–90–90 corresponde a que, em 2020, 90% das pessoas que vivem com VIH sejam diagnosticadas, 90% das pessoas diagnosticadas recebam TARV e 90% das pessoas que recebem TARV tenham as suas cargas virais indetectáveis (31).

Até ao momento, um arsenal de 31 medicamentos aprovados pela Food and Drug Administration (FDA) está disponível para o tratamento de infeções por VIH-1 (Tabela 1)

(33). Esses fármacos estão distribuídos em seis classes distintas com base em seus mecanismos de ação e perfis de resistência: (1) inibidores da transcriptase reversa análogos de nucleósidos (NRTIs), (2) inibidores da transcriptase reversa não-nucleósidos (NNRTIs), (3) inibidores da integrase (IN), (4) inibidores de protease (PIs), (5) inibidores de fusão e (6) antagonistas de co-receptores (Tabela 1) (33).

A terapia antirretroviral inicial, selecionada com base nos testes de resistência basal e nas características do paciente, continua baseada na combinação de 2 NRTIs e um terceiro agente, um IN, ou NNRTI ou um PI (32).

Os ARVs também são utilizados para prevenir a transmissão vertical do VIH, da mãe ao filho, para reduzir a transmissão a parceiros sexuais serodiscordantes e para prevenir a aquisição do vírus por pessoas que não estão infectadas, mas que podem ter sido expostas ou apresentam comportamentos de risco para a aquisição da infecção (profilaxia pós-exposição (PEP) e Profilaxia Pré-Exposição (PrEP)) (31).

A transcriptase reversa é alvo de duas classes distintas de agentes antirretrovirais: os NRTIs, que são os análogos aos nucleósidos que constituem parte dos substratos de polimerização, e os não análogos aos nucleósidos (NNRTIs), que se ligam a uma bolsa alostérica não catalítica na enzima (Figura 8) (33). Juntos, os 15 agentes licenciados nessas duas classes são responsáveis por quase metade de todos os medicamentos antirretrovirais aprovados (Tabela 1) (33). Embora os NRTIs e NNRTIs difiram em relação ao seu local de interação na enzima e mecanismo molecular, ambos afetam a atividade de polimerização do DNA da enzima e bloqueiam a geração do DNA pró-viral ((Figura 8) (33).

Os PIs (inibidores de Protease) bloqueiam a proteólise da poliproteína viral, uma etapa necessária para a produção de partículas virais infecciosas, ao inibir o sítio ativo da protease (Figura 8) (33). Esta classe de fármacos, está entre os agentes mais potentes desenvolvidos até hoje, mas são compostos grandes, semelhantes a peptídeos, que geralmente exigem a co-administração de um agente “impulsionador” para inibir seu metabolismo e aumentar os níveis séricos do fármaco (33).

Todos os inibidores da integrase em desenvolvimento têm como alvo a reação de transferência da cadeia e são, portanto, referidos como INI ou, mais especificamente, inibidores da transferência da cadeia da integrase (InSTIs) (Figura 8) (33,35).

Os Inibidores de Fusão ou inibidores de entrada do VIH-1 podem ser subdivididos em classes distintas baseadas na rutura, inibição de alvos e passos distintos no processo (33).

Os inibidores de fusão (Tabela 1) e (Figura 8) (33) incluem apenas a molécula (T-20, enfuvirtide) com potente atividade antiviral *in vivo* (33).

Os antagonistas do CCR5 ligam-se a bolsas hidrofóbicas dentro das hélices transmembranares do CCR5 (33). Este local não se sobrepõe aos locais de ligação dos agonistas de CCR5 ou do envelope de VIH-1 (33). Em vez disso, a ligação ao fármaco induz e estabiliza uma conformação do receptor que não é reconhecida por nenhum dos dois (33). Assim, essas moléculas são consideradas inibidores alostéricos.

O composto MVC (maraviroc) foi aprovado para uso terapêutico pela FDA em 2007 (Tabela 1). O MVC liga-se a uma cavidade transmembrana hidrofóbica do CCR5, que altera a conformação da segunda alça extracelular do recetor e previne a interação com a alça do tronco V3 da gp120 (Figura 8) (33).

Classe	Mecanismo	Fármaco
Inibidores Nucleósidos da Transcriptase Reversa (NRTI)	INTR são mimetizadores de nucleótidos / nucleosídeos que são incorporados na cadeia nascente de DNA, inibindo a atividade da transcriptase reversa na fase inicial do ciclo	Zidovudina (AZT)
		Didanosina (ddI)
		Zalcitabina (ddC)*
		Estavudina (d4T)
		Lamivudina (3TC)
		Abacavir (ABC)
		Tenofovir (TDF)
		Emtricitabina (FTC)
Inibidores de Protease (IP)	IP são peptídeos miméticos que se ligam covalentemente no sítio ativo da protease viral, impedindo a maturação viral na etapa final do ciclo	Saquinavir (SQV)
		Ritonavir (RTV)
		Indinavir (IDV)
		Nelfinavir (NFV)
		Amprenavir (APV)
		Lopinavir (LPV/r)
		Atazanavir (ATV)
		Fosamprenavir (fAPV)
		Tipranavir (TPV)
		Darunavir (DRV)
Inibidores Não Nucleósidos da Transcriptase Reversa (NNRTI)	INNTR foram projetados para se ligarem bolsa alostérica da TR, modificando a estrutura da enzima e desativando o sítio ativo no domínio da polimerase	Nevirapina (NVP)
		Delavirdina (DLV)*
		Efavirenz (EFV)
		Etravirina (ETR)
		Rilpivirina (RPV)
		Doravirina (DOR)

Inibidor de Fusão (IF)	<p>IF são pequenos peptídeos que se ligam ao domínio hidrofóbico da gp41, evitando a fusão das membranas viral e celular</p>	<p>Enfurvirtida (T-20)</p>
Antagonista de CCR5 (AC)	<p>AC são pequenas moléculas que se ligam a CCR5, modificando a sua estrutura e bloqueando sua interação com gp120</p>	<p>Maraviroc (MVC)</p>
Inibidores de Integrase (IN/ InSTIs)	<p>IN ligam-se à integrase viral e bloqueiam a integração do DNAc no genoma da célula hospedeira</p>	<p>Raltegravir (RAL) Dolutegravir (DGT) Raltegravir (RAL)</p>

Tabela 1: Resumo dos fármacos antirretrovirais

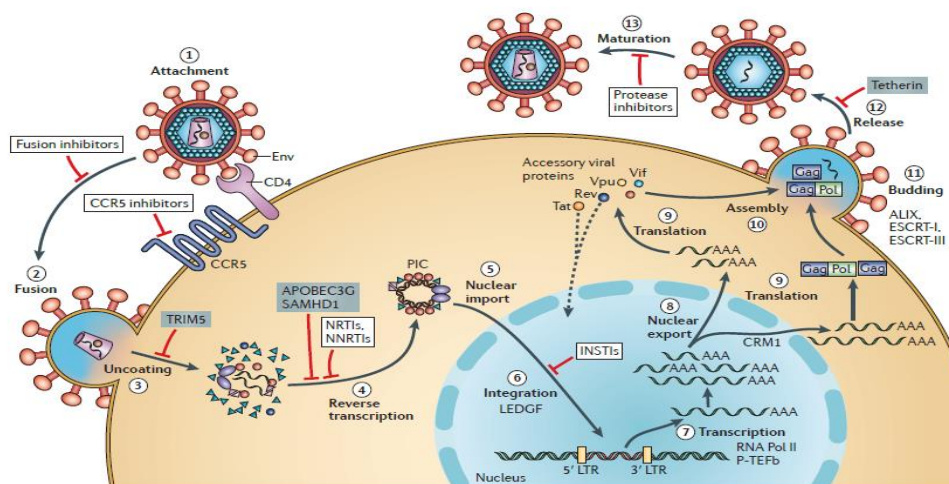


Figura 8: Visão geral do ciclo de replicação do VIH-1 e os mecanismos de ação dos ARVs.

Fonte: ENGELMANN; CHEREPANOV, 2012

1.8 Mecanismos de Resistência aos ARVs

Com o advento da terapia antirretroviral acreditou-se que a infecção pelo VIH poderia ser parada ou desacelerada. Contudo, nenhum fármaco ou combinações de fármacos ARVs estudados conseguem parar completamente a replicação do vírus (36). Por isso, todos os fármacos antirretrovirais atuais, incluindo as classes mais recentes, correm o risco de se tornar parcial ou totalmente inativos devido ao surgimento de vírus resistentes a estes medicamentos (34).

A resistência do VIH aos antirretrovirais pode vir a tornar-se num problema grave, uma vez que esta tem vindo a aumentar, sendo que nos últimos anos tem aumentado sobretudo em países de baixa e média renda.

A resistência aos antirretrovirais para o VIH (HIVDR) é um fenómeno evolutivo causado por uma alteração (mutação) no genoma do VIH que afeta a capacidade de um determinado fármaco ou combinação de fármacos em bloquear a replicação do vírus, favorecendo a seleção de estirpes virais melhor adaptadas para sobreviver no ambiente sob pressão seletiva de fármacos (34,37).

O fenómeno de resistência pode ser classificado em resistência primária, quando surge num doente em que não há história de terapia antirretroviral (TARV) prévia, ou resistência secundária, quando esta se desenvolve em doentes que foram expostos ao (s) medicamento (s) ARV(s) (37). A resistência secundária é comum quando a TARV não consegue suprimir totalmente a replicação do RNA do VIH no plasma do hospedeiro (37).

Existem alguns mecanismos de resistência aos ARVs já identificados e que se revestem de importância para estudos epidemiológicos na saúde pública.

A resistência aos NRTIs é mediada por dois mecanismos: pirofosforólise dependente de ATP, que é a remoção dos NRTIs da extremidade 3' da cadeia nascente, reversão da terminação da cadeia, e o aumento da discriminação entre o substrato de desoxirribonucleído nativo e o inibidor (33). E um segundo mecanismo de resistência que é a prevenção da incorporação dos NRTIs na cadeia nascente (33).

As mutações que ocorrem nos inibidores nucleósidos da transcriptase reversa são classificadas como mutações associadas a nucleósidos (NAMs) ou mutações análogas da timidina (TAMs) (33).

As TAMs promovem pirofosforólise e estão envolvidos na excisão da Zidovudina (AZT) e Stavudina (d4T). As alterações dos aminoácidos TAM na transcriptase reversa do VIH-1 incluem duas vias distintas: a via TAM1 (M41L, L210W, T215Y e ocasionalmente D67N) e a via TAM2 (D67N, K70R, T215F e 219E / Q) (33).

As mutações associadas a inibição da incorporação dos NRTIs na cadeia nascente, incluem o M184V / I e o K65R.

Muitas mutações primárias e secundárias aos NRTIs (ou combinações destas) demonstraram diminuir a função da transcriptase reversa e a capacidade de replicação viral (33). Embora vários estudos tenham sugerido um potencial de benefício clínico associado à redução da capacidade replicativa de variantes resistentes aos NRTIs, é importante observar que mutações adicionais podem se acumular na presença de tratamento contínuo, resultando em níveis mais altos de resistência (33).

A perda na adequação replicativa devido a mutações de resistência a fármacos (na ausência de fármacos) também pode ser compensada pela acumulação de mutações secundárias (33).

Em contraste com as reduções significativas na aptidão replicativa observadas com resistência a outras classes de fármacos, com NNRTIs, alterações de um único nucleotídeo podem resultar em resistência de alto nível (33).

Uma menor barreira genética, um impacto mínimo no condicionamento reprodutivo e a lenta reversão dessas mutações em pacientes na ausência de fármacos contribuem para a transmissão e estabilidade do VIH-1 resistente a NNRTI na população (33).

Em relação aos inibidores de integrase ou InSTIs, as mutações que geram resistência a esta classe de fármacos, quase sempre afetam as sequências de aminoácidos do sítio ativo da enzima, próximo aos resíduos de aminoácidos que coordenam os cofatores de magnésio essenciais (33).

Os inibidores de protease (IPs), devido ao seu papel vital no ciclo de vida do VIH-1 e o tamanho relativamente pequeno (11 kDa), inicialmente era esperado que a resistência a estes fosse rara (33). No entanto, a protease possui grande plasticidade, com polimorfismos observados em 49 dos 99 codões e mais de 20 substituições conhecidas por estarem associadas à resistência (33).

Todos os PIs compartilham estruturas químicas relativamente similares e a resistência cruzada é comumente observada (33). Para a maioria dos PIs, as mutações de resistência primária agrupam-se nas proteínas próximo do sítio ativo da enzima (sítio alostético), em posições localizadas no local de ligação do substrato / inibidor (33). Estas alterações de aminoácidos geralmente têm um efeito deletério no fitness viral (33).

Para além das mutações no gene da protease, as alterações localizadas nos oito principais locais de clivagem por protease - isto é, genes *gag* e *pol* - foram associadas à resistência aos inibidores da protease (33). As mutações nos locais de clivagem são melhores substratos para a protease e, assim, compensam parcialmente a perda de aptidão viral associada à resistência (33).

Em relação aos inibidores de fusão, demonstrou-se que a resistência aos inibidores alfa-helicoidais iniciais é mediada por mutações na região de repetição do heptad amino-terminal da gp41 (33).

A resistência aos antagonistas do CCR5, é avaliada através de estudos de resistência aos MVC, única droga desta classe desenvolvida e aprovada pela FDA, cujo mecanismo ação é a ligação desta a uma proteína da célula hospedeira (33). Assume-se que a resistência ao MVC dar-se pela fenotipagem do vírus do paciente, visto que as populações deste indivíduo pode ser maioritariamente X4 trópico, donde este fenótipo viral adentra a célula pela ligação ao coreceptor CXCR4, em detrimento ao CCR5 cuja a forma de resistência é diferente dos outros ARVs (33).

1.9 Caracterização sociodemográfica de São Tomé e Príncipe (STP)

São Tomé e Príncipe é um estado insular composto por duas ilhas e algumas ilhotas, localizado no golfo da Guiné, com uma área geográfica de 1001 km², e cerca 200 mil habitantes (Figura 9) (38). As ilhas estiveram desabitadas até 1470, quando os navegadores portugueses João de Santarém e Pedro Escobar as descobriram, e foi então uma colónia portuguesa desde o século XV até sua independência em 12 de Julho de 1975 (39).

Administrativamente, as ilhas encontram-se divididas em seis distritos, todos situados na ilha de S. Tomé (Agua Grande, Mé-zochi, Lobata, Lembá, Caué, Cantagalo) e a Região Autónoma do Príncipe (Figura 9) (39,40).

São Tomé e Príncipe tem um território e uma população de dimensão reduzida, e possui uma economia fragilizada, que lhe coloca na posição de recetor líquido de ajuda pública para o desenvolvimento (39). Com isto, o país apresenta uma economia das mais pequenas de África, com o seu PIB representar apenas 0,012% do PIB de África, facto que lhe coloca no 49.º lugar no ranking Banco Africano de Desenvolvimento (41). O seu PIB per capita em 2018 foi de USD 1,697 segundo o Banco Mundial Africano (42).

No que respeita à demografia, a população são-tomense é bastante jovem, e tem crescido consideravelmente nas últimas décadas, com mais da metade (53,6%) com menos de 35 anos de idade (39). Além disso, mais de 90% da população ativa possui algum nível de instrução (nível básico (48%), nível secundário (41%) e nível profissional/ superior (4%)) (39).

O Índice de Desenvolvimento Humano (IDH) do país foi de 0.589 em 2018, mantendo-se assim no 143º lugar do ranking do Programa das Nações Unidas para o Desenvolvimento (PNUD) (40,43). Além disso, a esperança média de vida da população geral é de 67 anos, e a taxa de mortalidade infantil é de 51 por 1000 nados-vivos (40,43). O acesso a uma fonte melhorada de água abrange 97% da população e 60% da população possui eletricidade (40).

A pobreza atinge cerca de 53,3% da população, e isto tem implicações negativas para os sectores económico, político, social e institucional (39).



Figura 9: Mapa de São Tomé e Príncipe. Fonte: <https://www.telanon.info>

1.10 VIH em São Tomé e Príncipe

Os dados disponíveis indicam que o país apresenta uma epidemia fraca e concentrada (44). Contudo, teme-se que haja um aumento da transmissão da infeção nas próximas décadas, dadas as fragilidades económicas e sociais existentes no país (45).

Segundo o Inquérito Demográfico Sanitário (IDS) de 2008, a prevalência do VIH na população são-tomense era de 1,5% em adultos dos 15 aos 49 anos de idade, sendo que os homens apresentavam uma prevalência de 1,7% comparativamente com as mulheres (1,3%) (44). Por outro lado, segundo os dados do MICS (Inquérito de Agrupamentos de Indicadores Múltiplos), a prevalência era de 0,5% em 2014 (46).

A taxa de cobertura da TARV em 2017 foi de 78,2%, e a taxa de abandono foi de 0,8% no mesmo ano (47). A taxa de sobrevivência aos 12 meses após o TARV foi de 76,3%, em 2016 (47). Dadas as vulnerabilidades sociais, económicas, políticas e geográficas de São Tomé e Príncipe, é importante existir uma atenção redobrada para as doenças sexualmente transmissíveis, como o VIH/SIDA, para se evitar o incremento da infeção numa população

tão reduzida, mas estrategicamente um possível foco de transmissão para outras partes do mundo.

Por outro lado, este pequeno arquipélago faz fronteira marítima com alguns países da costa Africana, considerados endêmicos para o VIH, como os Camarões (3,7% de adultos -15-49 anos- infetados), a Guiné-Equatorial (6,5% de adultos infetados), o Gabão (4,2% de adultos infetados), e a Nigéria (2,8% de adultos infetados) (48), todos com registo de prevalências altas da infeção pelo vírus em adultos jovens.

O primeiro caso da infeção pelo vírus no país foi reportado em 1987, e notificado em 1990 no distrito de Água-Grande (45).

Dados preliminares disponíveis no primeiro estudo de epidemiologia molecular do VIH-1 realizado no país, no período de 2004 a 2007, com amostras de doentes seropositivos, apontou para a presença de estirpes com uma grande variedade genética e com elevada percentagem de vírus recombinantes com padrões de recombinação distintos, sugerindo múltiplas introduções no país (45).

No entanto, ainda pouco se conhece sobre a epidemiologia molecular do VIH-1 em S.Tomé e Príncipe, as características e determinantes sociais que influenciam a rede de transmissão do vírus, e a origem do mesmo no país. E estando S.Tomé e Príncipe localizado numa região considerada endêmica, associada às suas vulnerabilidade, este estudo propõe utilizar métodos de filogenética e filogeografia para esclarecer a origem temporal e geográfica do vírus no país, caracterizar o seu padrão de dispersão e identificar mutações de resistência aos antiretrovirais que circulem no país.

2. Objetivos

2.1 Geral

Reconstruir a história da epidemia do VIH-1 em São Tomé e Príncipe, incluindo a caracterização das estirpes circulantes, a origem temporal e geográfica da epidemia.

2.2 Específicos

- 1) Classificar estirpes de VIH-1 circulantes em doentes infetados por VIH-1 residentes em STP.
- 2) Determinar a origem temporal e geográfica da epidemia e as formas de entrada e disseminação do vírus no arquipélago de STP.
- 3) Identificar e classificar as mutações de resistência aos antirretrovirais que circulam em doentes infetados por VIH-1 residentes no arquipélago de STP.

3. Material e Métodos

3.1 Casuística

Este trabalho de investigação tratou-se de um estudo observacional transversal baseado em dados secundários recolhidos no âmbito da tese de doutoramento da investigadora Idalina Bonfim, e que estavam disponíveis na base de dados do hospital de Egas Moniz, onde foram geradas as sequências. Este dataset incluiu 93 sequências do gene *pol* (polimerase) do VIH-1 correspondentes a protease (1-99) completa e dois terços da transcriptase reversa (1-335), anotadas em formato FASTA, com um comprimento de 1302 bp. As amostras foram colhidas no período de 2004 a 2007 em pacientes seropositivos seguidos em várias instituições de STP.

Dentre estas 93 sequências *pol* do VIH-1, 58 (62%) eram de pessoas não tratadas com os antirretrovirais (ARVs) (naives), e 35 (38%) eram de pessoas tratadas com antirretrovirais, no âmbito do Programa Nacional de Luta contra a SIDA em STP. Trinta e um (33%) dos participantes eram do sexo masculino, cinquenta (54%) do sexo feminino e doze (13%) não possuíam informação relativa ao género. A média das idades da população do estudo foi de 34 anos, sendo a mínima de 7 anos e a máxima de 63 anos, de ambos os sexos (45).

Estas amostras foram submetidas a genotipagem, a análise filogenética e filogeográfica, e a identificação de mutações de resistência aos antirretrovirais, com recurso aos programas bioinformáticos.

3.2 Determinação dos subtipos do VIH-1

As sequências nucleotídicas (n=93) do VIH-1 foram submetidas a alinhamento múltiplo de sequências e subsequentemente editadas utilizando os programas Bioedit e Mega 7. Foram eliminadas as sequências que estavam em duplicado, totalizando-se no final 88 sequências. Estas sequências foram renomeadas/recodificadas, com recurso ao software Excel e ao software Fabox (<http://users-birc.au.dk/palle/php/fabox/index.php>).

Após a atribuição de um novo código, as sequências nucleotídicas (n=88) foram analisadas com algoritmos bioinformáticos de subtipagem de sequências genômicas de VIH-1: *SCUEAL*, (http://classic.datamoney.org/dataupload_scueal.php), *COMET* (<https://comet.lih.lu/>), e *REGA HIV-1 SubtypingTool* versão 3.0 (<http://dbpartners.stanford.edu:8080/RegaSubtyping/stanford-hiv/typingtool>). Estas ferramentas bioinformáticas são baseadas em algoritmos em que as sequências *pol* são alinhadas com as de referência dos subtipos puros e das Formas Circulantes Recombinantes do grupo M do VIH-1 (45). Em situações de discrepâncias entre os algoritmos (*COMET*, *SCUEAL* e *REGA*), optou-se pela classificação fornecida pelo *REGA Subtyping Tool*. Nos casos em que não foi possível a subtipagem por nenhum destes três algoritmos, as sequências foram analisadas separadamente com o programa de *jpHMM-HIV* da *University of Göttingen*.

3.3 Análise Filogenética

3.3.1 Alinhamento e construção da árvore (inferência de árvore filogenética)

Para a análise e construção de árvores filogenéticas, foram utilizadas apenas sequências classificadas como a forma circulante recombinante CRF 02_AG (n=27), que corresponderam ao subtipo mais prevalente (31%) em São Tomé, de acordo com as 88 sequências analisadas.

A estas sequências CRF02_AG do gene *pol* do VIH-1 colhidas em S.Tomé (n=27), foi atribuído o nome de “STP”. O mesmo processo foi feito para os outros conjuntos de sequências, tais como as sequências CRF02_AG portuguesas “PT” (n=150), e as dos imigrantes são-tomenses tratados em Portugal “STP_IM” (n=16), retirados da base de dados do hospital de Egas Moniz (VIH RegaDB). Na análise, foram também incluídas todas as sequências CRF 02_AG (n=19827) disponíveis na base de dados de Los Alamos (<https://www.hiv.lanl.gov>), definida como “WORLD”; que foram utilizadas como controlos, de forma a compreender a entrada e dinâmica de transmissão da epidemia VIH-1 em STP.

Esses conjuntos de sequências *pol* CRF02_AG do VIH-1 correspondentes a protease e a transcriptase reversa (RT/PR), foram todas combinadas e alinhadas juntamente com a

sequência de referência - HXB2 - do genoma do VIH-1 no programa de alinhamento rápido, Virulign Master.

Depois do alinhamento, as sequências foram editadas no Mega 7, tendo sido eliminadas sequências com gaps correspondentes a fragmentos superiores a 10% do tamanho do alinhamento (sequências curtas), e/ou aquelas compostas apenas ou por PR ou RT. O dataset assim constituído totalizou 5379 sequências CRF02_AG. Em seguida, deste alinhamento, foram excluídas as colunas do alinhamento correspondentes às principais mutações de resistência aos ARVs, gaps, sendo que o alinhamento ficou com um comprimento de 792 nucleótidos.

Em seguida construiu-se uma árvore filogenética de Máxima Verosimilhança / *Maximum Likelihood* (ML) no programa FastTree 2.1.10 com o alinhamento de 5379 sequências CRF02_AG de 792 nucleótidos, posteriormente visualizada no software FigTree v1.4.4, para análise e estudo do padrão de distribuição filogenético das mesmas. Os clusters monofiléticos formados com as sequências de São Tomé e Príncipe (STP e STP_IM) na árvore filogenética de ML, com uma pontuação de ramificação com alto suporte (aLRT ≥ 0.70), foram identificados e selecionados para a reconstrução evolutiva e demográfica espacial do CRF02_AG VIH-1.

3.3.2 Filogeografia (reconstrução evolutiva e demográfica)

Para a reconstrução evolutiva e espaço-temporal das sequências, converteu-se o ficheiro FASTA, contendo 187 sequências CRF02_AG para um ficheiro em formato XML, no programa Utilitário de Análise Evolucionária Bayesiana (BEAUti v.1.10.1) e que é lido pelo software BEASTv1.10.1. Este programa serve para definir os modelos evolutivo, de genética populacional, de relógio molecular, de filogeografia e as opções da análise Markov Chain Monte Carlo (MCMC) (<https://beast.community> (Drummond e Rambaut, 2007)).

Na segunda etapa, executou-se o BEAST v1.10.1 (programa Bayesiano de Análise Evolutiva por Árvores de Amostragem) utilizando o arquivo de entrada contendo os dados, o modelo e as configurações feitas no BEAUti v.1.10.1. De seguida, na terceira etapa, os programas

Tracer v1.7.1, LogCombiner v1.10.1, TreeAnnotator v1.10.1 e Fig.Tree v.1.4.3, serviram para análise, combinação, anotação e visualização das árvores geradas pelo BEAST v1.10.1, respetivamente. Após estas etapas procedeu-se a análise e interpretação dos dados gerados e visualização do padrão de dispersão filogenética do CRF02_AG pelo software SPREAD.

No BEAUTi v1.10.1, para a configuração dos dados selecionou-se o ficheiro BEAST XML contendo as 187 sequências VIH-1 CRF02_AG, e configurou-se os parâmetros de análises do modelo de substituição de nucleótidos HKY, com a heterogeneidade nas taxas de substituição entre os locais distribuída de acordo com uma distribuição gama e sem a partição dos parâmetros da taxa pela posição dos codões para as sequências. Para o ficheiro dos *traits* contínuos contendo as latitudes e as longitudes, definido como 'Location', escolheu-se o modelo de substituição Cauchy RRW. Para a configuração do modelo de relógio molecular optou-se pelo modelo relaxado não correlacionado distribuído de acordo com uma distribuição lognormal, e foi utilizado um modelo populacional Bayesian Skygrid, com 36 pontos de grade como número de parâmetros. No painel de estados definiu-se reconstruir os estados em todos os ancestrais apenas para os traits. Na definição da cadeia MCMC (Bayesiana Markov Chain Monte Carlo) configurou-se o comprimento desta para 1×10^8 , gravadas num ficheiro log a cada 10.000 etapas (26).

Após gerar uma amostra aleatória de árvores filogenéticas com raízes a partir de uma distribuição a priori de cada conjunto de dados baseado no modelo de coalescência Bayesiana Markov Chain Monte Carlo (MCMC) (49), no BEAST v1.10.1, procedeu-se a execução de três corridas desta MCMC de 1×10^8 etapas.

Para analisar a execução da MCMC do BEAST, utilizou-se o programa Tracer v.1.7.4, que permite produzir estimativas das distribuições marginais posteriores dos parâmetros do nosso modelo, acompanhando sempre o tamanho efetivo da amostra (ESS-effective sample sizes), que é uma medida do número de estimativas independentes dos parâmetros. Posteriormente utilizou-se o programa LogCombiner v1.10.1 para combinar as três árvores, assim como os três ficheiros log gerados em cada corrida. Para a sumarização das árvores, a árvore de máxima credibilidade de clade (MCC) foi calculada e anotada utilizando-se o TreeAnnotator v1.10.1 que permite anotar todas as informações resumidas de toda as gerações de árvores

numa única árvore, com um burn-in de 10% para as árvores, ou seja, os primeiros 10% de árvores foram descartadas. Posteriormente as relações filogenéticas e temporais com as amostras de S.Tomé e Príncipe foram visualizadas e analisadas no Fig.Tree v.1.4.3 e v.1.4.4. Para a reconstrução da dinâmica evolucionária espaço-temporal dos dados, o ficheiro do resumo das árvores geradas pelo TreeAnnotator v1.10.1 foi analisada e visualizada pelo software SpreaD3 v. 0.97 que permite a visualização em mapas personalizados (<https://beast.community>) gerados em páginas HTML.

3.4 Avaliação da resistência aos ARVS (stanfordDB 8.7)

Para análise e interpretação do perfil das mutações de resistência das sequências *pol* do VIH-1 em formato FASTA, utilizou-se o algoritmo de análise e interpretação da resistência aos ARVs da Universidade de Stanford/ *Stanford HIV Drug Resistance Database* (stanfordDB 8.7) (<https://hivdb.stanford.edu>), que calcula a suscetibilidade das estirpes às diferentes classes de fármacos NRTI, NNRTI, PI e / ou INSTI usando um sistema de pontuação para cada mutação (DRMs) de resistência a ARVs nos diferentes genes que são atingidos pelos ARVs (50). Para identificar as mutações de resistência, as sequências *pol* do DNA proviral são alinhadas e comparadas com as sequências do consenso B da protease e da transcriptase reversa (<http://hivdb.stanford.edu/>) (45).

Para cada estirpe, é feita a soma da pontuação das mutações identificadas (45). Para a classificação dos níveis de resistência, o Sistema de Interpretação do Teste de Resistência Genotípica do HIVDB (GRT) relata cinco diferentes níveis possíveis de resistência a cada ARV: suscetível, resistência potencial de baixo nível, resistência de baixo nível, resistência intermediária e resistência de alto nível (50).

A "resistência de baixo nível" é atribuída quando um vírus tem DRMs associados a uma redução da suscetibilidade a ARV *in vitro* ou a uma resposta virológica subótima ao tratamento com ARV. "Resistência intermediária" é atribuída quando, embora exista uma alta probabilidade de que a atividade de um ARV seja reduzida na presença de DRMs de um vírus, o ARV provavelmente ainda manterá uma atividade antiviral significativa contra o

vírus. "Resistência de alto nível" é atribuída quando um vírus possui DRMs previstos para conferir um nível de resistência semelhante ao observado em vírus com os mais altos níveis de suscetibilidade reduzida in vitro ou em vírus com pouca ou nenhuma resposta virológica ao tratamento ARV (50).

Fármacos com níveis de resistência acima de 3 foram considerados na classificação deste estudo, sendo que seguindo o padrão do sistema GTR, valores iguais a 3, foram considerados como resistência baixa, valores iguais a 4, como resistência moderada e valores iguais a 5, resistência alta.

4. Resultados

4.1 Genotipagem (subtipagem)

A determinação dos subtipos das 88 sequências do gene *pol* do VIH-1 dos indivíduos infetados, colhidas em S.Tomé e Príncipe no período de 2004 a 2007 e submetidas aos algoritmos bioinformáticos (*SCUEAL*, *REGA*, *COMET*), evidenciou uma elevada variabilidade genética do grupo M do VIH-1 em STP, com predomínio dos subtipos não-B (47%) e recombinantes (47%). O **CRF02_AG** foi a estirpe com maior prevalência, 31% (27/88), seguido pelo **subtipo A** 17% (15/88), o **subtipo G** 9% (8/88), e outros recombinantes 16% (14/88). Ficou evidente que outras variantes virais como o **subtipo C** 7% (6/88), **subtipo B** 5% (5/88) **subtipo F** 5% (4/88), **subtipo J** 5% (4/88), **subtipo D** 2% (2/88), e o **subtipo H** 2% (2/88) também co-circulam no país. Para 1% das sequências (1/88) não foi possível a classificação de subtipo. Na Figura 10 é possível ver a distribuição de estirpes encontradas.

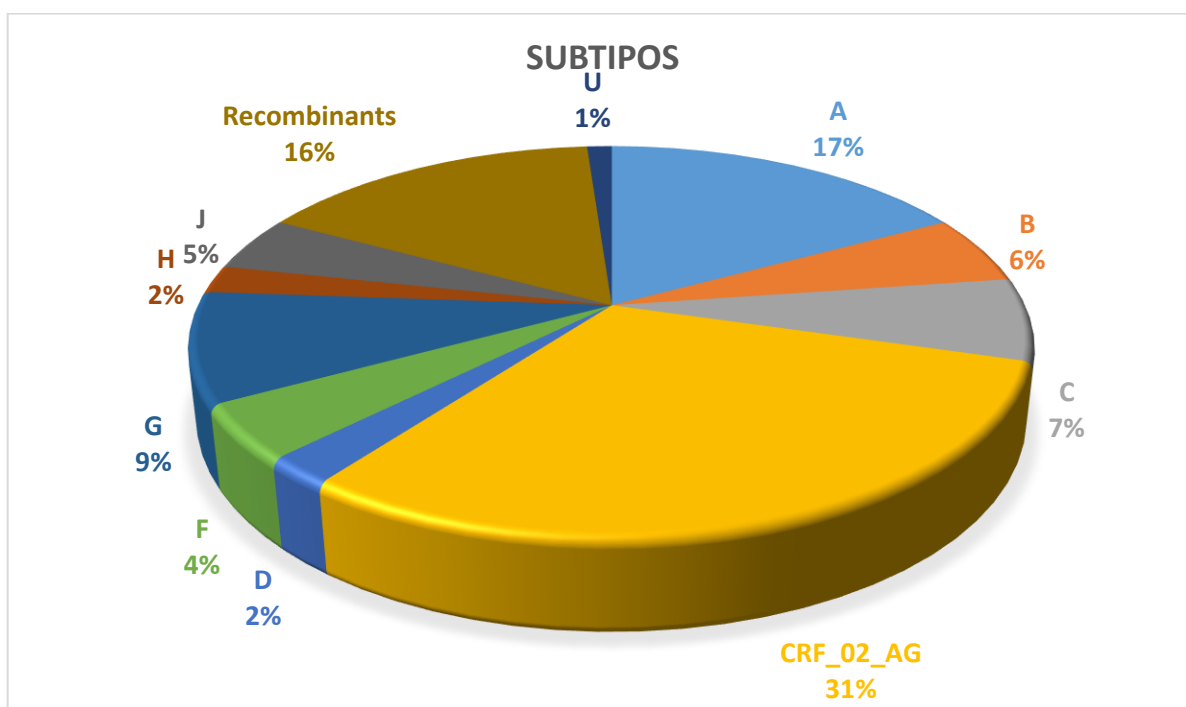


Figura 10: Distribuição dos subtipos do VIH-1, CRFs e outros recombinantes circulantes em S.Tomé e Príncipe

4.2 Análise Filogenética e Filogeográfica das sequências

A origem geográfica e temporal das estirpes CFR02_AG do VIH-1 de STP foram inferidas a partir da reconstrução filogenética de uma árvore MCC (*Maximum Credibility Clade*) das sequências globais *pol* do CRF 02_AG VIH-1 (n=187) usando a estrutura filogeográfica bayesiana inserida no pacote do BEAST.

Os resultados mostraram que a taxa evolutiva média das sequências *pol* do CRF02_AG VIH-1 do conjunto de dados, estimada numa escala de tempo cronológica foi de 3.6×10^{-3} com intervalo 95% HPD (Densidade Posterior mais alta) de $[2.0 \times 10^{-3} - 4.4 \times 10^{-3}]$ substituições / nucleótido / ano. A média da idade da raiz do ancestral comum mais recente do CRF02_AG foi de 44 anos com um intervalo 95% HPD [34.4 - 58.8], que remonta aos anos 1970.

A análise da árvore MCC revelou que o CRF02_AG do Gana é o ancestral comum mais recente de todas as sequências CFR02_AG do conjunto de dados, com uma PSP=1 (posterior state probability) (Figura 11). Este por sua vez é o ancestral comum mais recente das estirpes CRF02_AG introduzidas em S.Tomé e Príncipe, aproximadamente em 1984 (idade do nó=36; HPD: 1981 - 1988) (Figura 11) e (Figura 12).

Na análise filogenética da árvore MCC é possível observar-se que a partir da raiz formam-se dois ramos que dão origem a dois clusters, sendo que um dos ramos forma um pequeno cluster com apenas duas sequências (idade do nó=25 anos) e o outro forma um grande cluster (idade do nó=41 anos), de onde divergem várias sequências da amostra, e que dão origem a novos clusters, a partir da qual se forma uma grande rede de dispersão e introdução do CFR02_AG para outros países do globo (Figura 11).

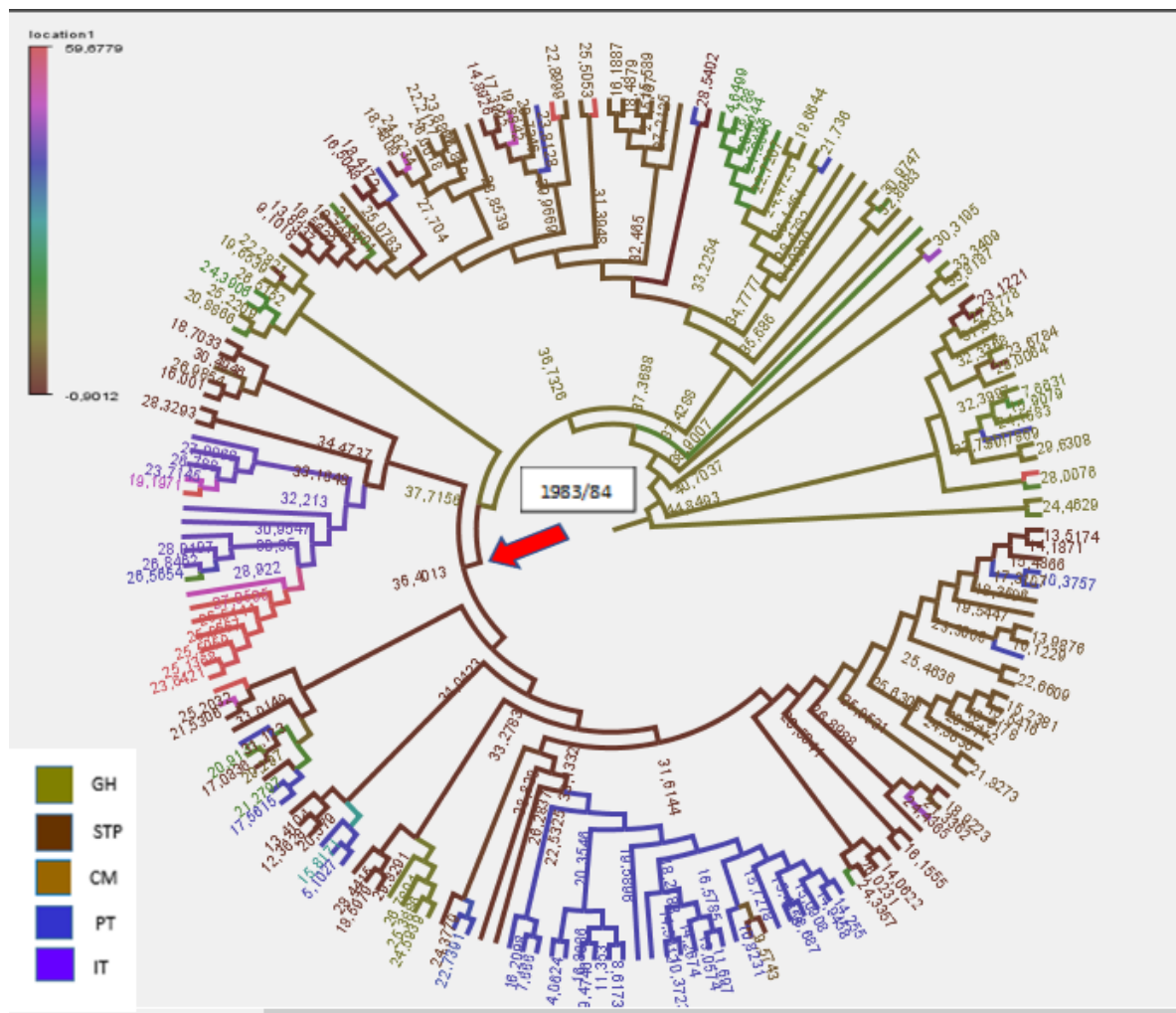


Figura 11: Escala temporal da árvore de credibilidade máxima clade bayesiana (MCC) das estirpes VIH-1 CRF02_AG são-tomenses e sequências globais próximas, extraídas do NCBI (n=187). As ramificações estão coloridas de acordo com o estado de localização mais provável do ancestral comum desse nó, conforme indicado na legenda. A seta vermelha indica o nó do primeiro ancestral das estirpes CRF02_AG São-Tomenses. Alguns países representados são Gana (GH), São Tomé e Príncipe (STP), Camarões (CM), Portugal (PT) e Itália (IT). A cor verde representa os clusters do Gana; a cor castanha representa os clusters São-Tomenses; o castanho-esverdeado representa os clusters Camaroneses; a cor azul representa os clusters Portugueses e o anil representa os clusters da Itália. A seta em vermelho indica a data aproximada da entrada do VIH-1 CRF02_AG.

Current date: 1983/08/28

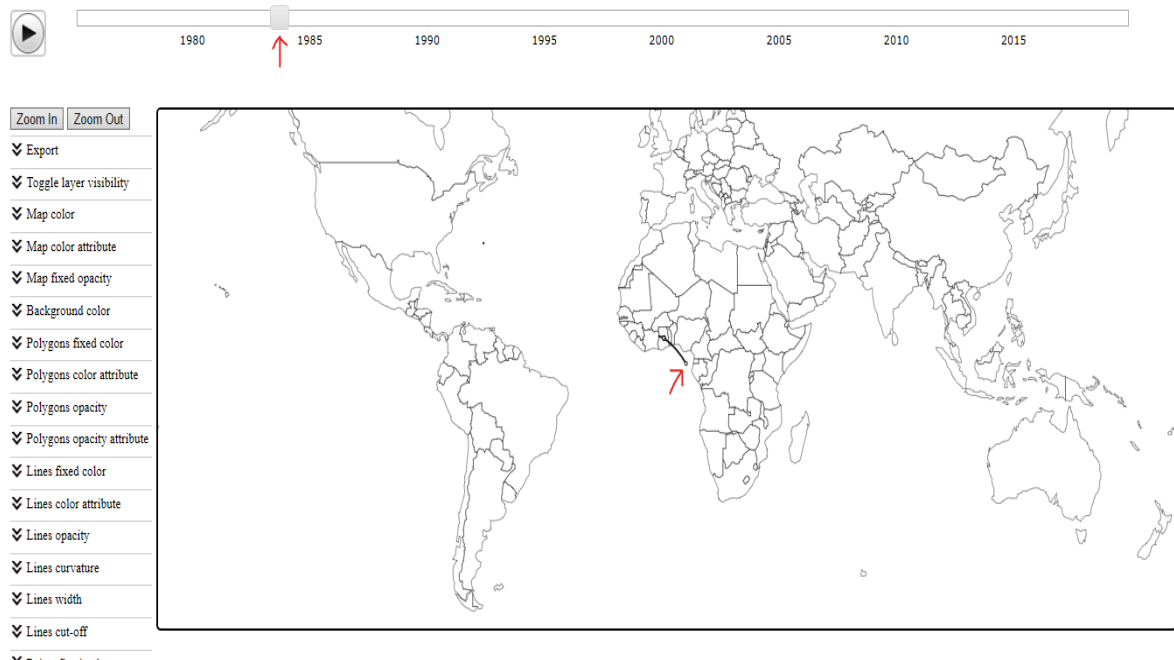


Figura 12: Mapa ilustrativo da entrada da estirpe VIH-1 CRF02_AG em São Tomé e Príncipe. Imagem gerada com recurso ao software SPREAD. As setas em vermelho indicam a provável data de entrada do CRF02_AG em STP.

Verifica-se na árvore uma grande dispersão das sequências de STP, que não se encontram agrupadas num só cluster (Figura 11). As sequências de STP formam pequenos sub-clusters dentro de clusters formados por sequências portuguesas e de clusters formados por sequências Camaronesas (Figura 11). Outras sequências encontram-se dispersas nas árvores formando pequenos clusters de duas a três sequências, com uma $PSP \geq 0.60$ (*posterior state probability*), ou formam pequenos clusters monofiléticos com sequências provenientes do Gabão, Gana e Benin nomeadamente.

No cluster das sequências Portuguesas, é possível constatar que duas sequências (STP_765003_2013 e STP_395292_2009) pertencentes aos imigrantes são-tomenses em Portugal, formam um cluster monofilético altamente suportado ($PSP=0.8$) com uma idade do nó a rondar os 9 a 10 anos atrás (2009-2010), sugerindo que a infeção destes dois indivíduos

aconteceu em Portugal, e que estão provavelmente relacionados epidemiologicamente (Figura 13). A topologia da árvore e as estimativas de localizações dos ancestrais comuns dos diferentes nodos indicam ainda que a origem da maior parte das estirpes de CRF02_AG introduzidas em Portugal terá sido a partir de estirpes provenientes de São Tomé e Príncipe (PSP=0.9) (Figura 11). Contudo sabe-se que existe um viés de amostragem, uma vez que se tem muitas sequências portuguesas e de S.Tomé e Príncipe, e uma escassez de estirpes dos outros países da CPLP que também são endêmicos para o VIH-1 CRF02_AG, países com os quais Portugal tem uma estreita relação de cooperação cultural, social e económica, não se podendo inferir desta forma que a primeira introdução deste recombinante tenha partido de S.Tomé.



Figura 13: Escala temporal da árvore de credibilidade máxima clade bayesiana (MCC) das estirpes VIH-1 CRF02_AG são-tomeneses e sequências globais próximas, extraídas do NCBI (n=187). Representação do cluster das sequências Portuguesas, com ramificações em azul. A barra em amarelo indica o sub-cluster formado entre as sequências de São-Tomé (em vermelho) com as sequências de Portugal, (PSP≈0.8).

4.3 Análise das Resistências aos ARVs

A submissão das sequências *pol* (n=88) do VIH-1 em formato FASTA no algoritmo da *Stanford University HIValg Program* (<https://hivdb.stanford.edu>), permitiu estimar a prevalência das mutações de resistência aos antirretrovirais existentes na população do estudo, e identificar o perfil das mutações associadas a resistência a protease e a transcriptase reversa.

Os resultados obtidos nesta análise evidenciaram que 18% (16/88) dos indivíduos apresentavam mutações de resistência aos antirretrovirais. Além de que, das 88 sequências analisadas, 15% (13/88) apresentavam mutações de resistência aos NRTI, 16% (14/88) apresentavam mutações de resistência NNRTI, e 2% (2/88) apresentavam mutações de resistência aos PI (Figura 14) e (Tabela 3).



Figura 14: Prevalência de resistências aos ARVs em estirpes de VIH-1 isoladas em S.Tomé e Príncipe

As mutações de resistência identificadas para cada classe de antirretrovirais foram as seguintes: **NNRTIs** (K103N/S, V106I/M, V179D/E, G190A, Y181C, A98G, E138Q); **NRTIs**: (M184V, D67E/N, T69D, K70R/N, M41L, K219Q, L210LW, T215Y); **PIs**: (V32I, M46I, I47V, I54L, V82T, I84V, F53L) (Tabela 2) e (Tabela 3).

Frequência das mutações de resistência								
NRTI			NNRTI			PI		
M184V	7	8%	V106I/M	5	6%	V32I	1	1%
D67E/N	3	3%	V179D/E	2	2%	M46I	1	1%
T69D	1	1%	G190A	2	2%	I47V	1	1%
K70R/N	2	2%	K103N/S	6	7%	I54L	1	1%
M41L	2	2%	A98G	1	1%	V82T	1	1%
K219Q	1	1%	E138Q	1	1%	I84V	1	1%
L210W	1	1%	Y181C	2	2%	F53L	1	1%
T215Y	2	2%						

Tabela 2: Tabela de distribuição das frequências relativas e absolutas das mutações de resistência consoante as classes de ARVs (N=88).

Também existiram indivíduos que apresentaram resistência a mais do que uma classe de ARVs. Dos 88 indivíduos da amostra, 8% (7/88) apresentaram resistência a uma única classe de ARVs, 9% (8/88) apresentaram resistência a duas classes de ARVs (NRTI+ NNRTI) e 1% (1/88) apresentaram resistência tripla a três classes de ARVs (NRTI+ NNRTI+ PI) (Figura 15).



Figura 15: Distribuição dos padrões de resistência nas três classes de fármacos ARVs.

Analizou-se também a frequência das mutações de resistência nos diferentes subtipos identificados nos indivíduos do estudo, e verificou-se que os subtipos F (75%) e B (60%) são os que apresentaram maior percentagem de mutações de resistência, seguidos pelo subtipo D com 50%, subtipo G com 38%, e o subtipo C, com 17%. As estirpes recombinantes (CRFs e URFs) foram as que apresentaram menores percentagens de mutações de resistência, sendo o CRF02_AG, a estirpe mais prevalente na nossa amostra, com apenas 7% das mutações de resistência. (Figura 16).

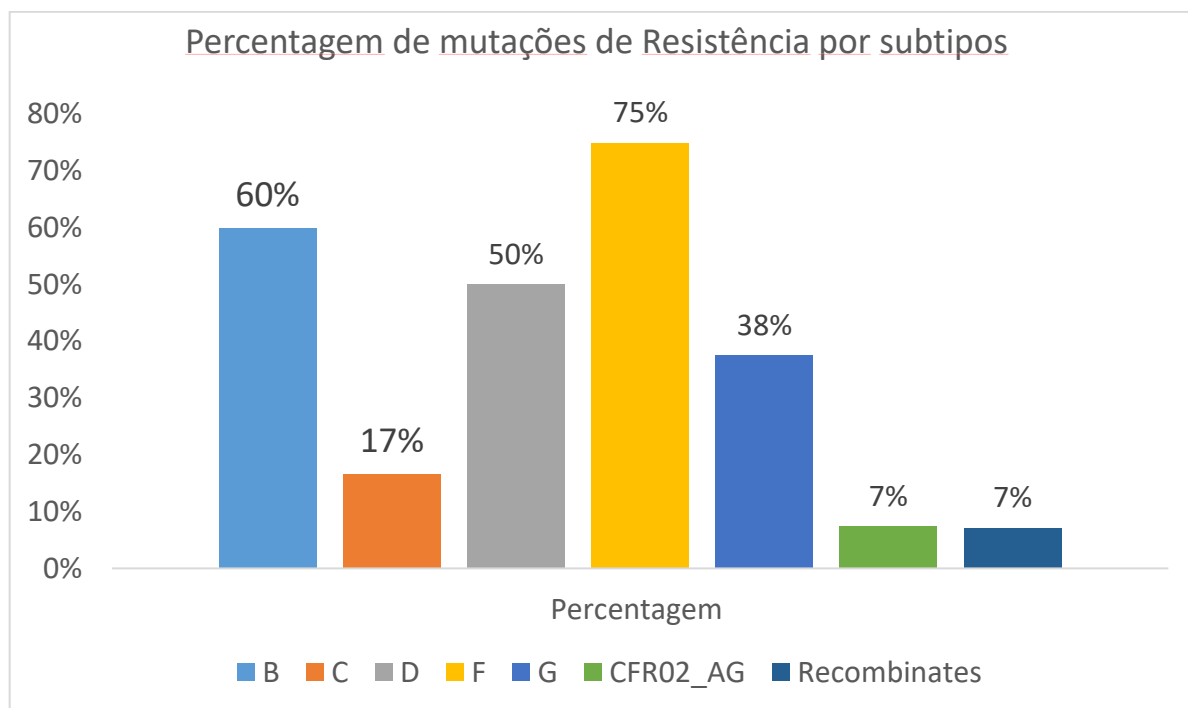


Figura 16: Frequência das mutações de resistência consoante as estirpes do VIH-1 identificadas

O gráfico abaixo mostra que os NRTIs e NNRTIs são as classes de ARVs que apresentam maior percentagem de resistência de nível alto (NRTI=10%; NNRTI=10), e moderado (NRTI=10% e NNRTI=7%), comparativamente aos PI (Figura 17).

Percentagem dos níveis de resistência por classe de fármacos

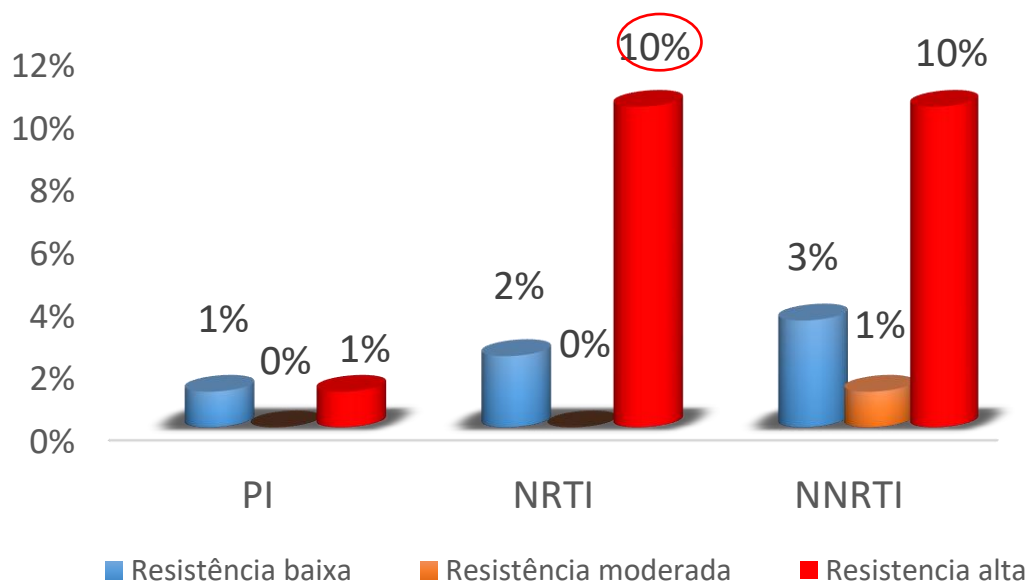


Figura 17: Distribuição das percentagens dos níveis de resistência consoante as classes de fármacos ARVs.

Os resultados obtidos em relação aos níveis de resistência a fármacos utilizados no tratamento primeira linha dos adultos com VIH em STP, demonstraram que Zidovudina, Lamivudina, Emtricitabina e Abacavir dos NRTIs, apresentaram maior percentagem de indivíduos com níveis de resistência moderados e altos. Para os NNRTIs, a Nevirapina, o Efavirenz e a Rilpivirina também apresentaram mais indivíduos com níveis de resistência moderados e altos. E para os PIs, verificou-se que os três antirretrovirais considerados como tratamento de primeira linha (Lopinavir, Darunavir, Atazanavir) apresentaram todos 1% de resistência de nível alto. Contudo estes três últimos ARVs da classe dos inibidores de protease (PI) eram usados como esquemas de segunda e terceira linha para os adultos (Lopinavir, Darunavir), e de primeira linha (Lopinavir) para as crianças menores de cinco anos, inclusive. O gráfico abaixo ilustra os resultados obtidos (Figura 18).

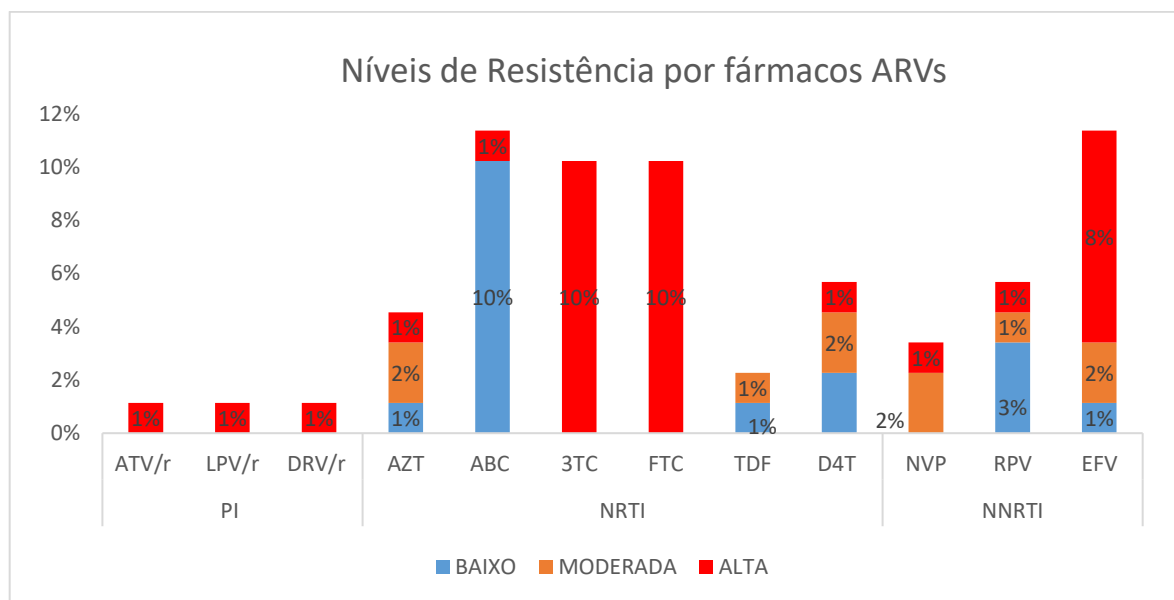


Figura 18: Percentagem dos níveis de resistência relativamente aos antirretrovirais usados no tratamento do VIH-1 em S.Tomé e Príncipe.

Sequências	Subtipos	Classe das drogas			
		IP Maior	IP Acessória	NRTI	NNRTI
STP06	D	-	-	K219Q	V106I
STP07	B	-	-	M41L,D67N,M184V, L210W, T215Y	V106I
STP09	C	-	-	-	K103N/S
STP10	B	-	-	-	K103N
STP17	B	-	-	K70N	V106I
STP21	B	V32I; M46I; I47V; I54L; V82T; I84V	L33F; Q58E	D67N; T69D; K70R; M184V	V106M; V179D; G190A
STP23	Recombinantes	-	-	M184V	-
STP29	CRF_02_AG	-	-	M184V	A98AG;V106I
STP34	F	-	-	M184V	K103N
STP35	U	-	-	M184I	E138Q;Y181C
STP36	G	-	-	M41L	K103N
STP38	G	-	-	D67E	V179E
STP41	F	-	-	M184V	Y181C
STP54	F	-	-	M184V	K103N/S; G190A
STP55	G	-	F53FL	-	-
STP88	CRF_02_AG	-	-	M184V,T215Y	K103N

Tabela 3: Resumo das sequências, subtipos e mutações de resistência identificadas

5. Discussão

5.1 Genotipagem

A ampla sequenciação do genoma do VIH-1 tem permitido a determinação dos subtipos, o estudo da evolução genética, epidemiologia molecular, sua relação com a patogénese da doença e resistência aos antirretrovirais, (51). Esta ferramenta tem possibilitado a monitorização das mudanças geográficas na epidemia mundial da SIDA, não obstante a elevada variabilidade genética do VIH (9).

A grande capacidade replicativa, recombinativa e evolutiva do VIH-1, assim como a existência de diferenças significativas na sua patogénese ou transmissibilidade (9) têm constituído um obstáculo para o desenvolvimento de vacinas que sejam realmente eficazes contra o vírus ou medicamentos capazes de o eliminá-lo. Por isso, a cura para a infeção VIH/SIDA e a travagem do avanço desta epidemia ainda é um desafio global, apesar dos imensos esforços e investimentos que foram feitos.

Este trabalho de investigação procura reforçar o conhecimento sobre a epidemiologia molecular do VIH-1 em S.Tomé e Príncipe, que é praticamente inexistente. Existe apenas um estudo desta natureza, que foi realizado entre 2004 e 2007 por Bonfim *et al* e que envolve amostras colhidas em doentes VIH-1 de STP entre 2004 e 2007 (45). Este estudo recicla essas sequências de STP para realizar uma análise mais completa e utilizando métodos mais avançados para analisar a epidemiologia molecular do VIH-1 em STP, com particular ênfase sobre o CRF02_AG que é a estirpe mais prevalente. Neste estudo utilizamos métodos de filogeografia para estimar a origem geográfica e temporal desta estirpe no arquipélago.

Baseado nas sequências do gene *pol* do VIH-1 submetidas a genotipagem pelos algoritmos bioinformáticos descritos na metodologia do trabalho, os resultados confirmaram uma grande variabilidade genética viral do grupo M a circular em S.Tomé e Príncipe, tal como o estudo de Bonfim *et al* já tinha encontrado (52). Os métodos de subtipagem utilizados neste estudo, mais robustos e desenvolvidos posteriormente ao trabalho de Bonfim *et al*, permitiram confirmar os seus resultados: a circulação de pelo menos sete subtipos e mais de uma forma recombinante em STP, com predomínio de subtipos não-B (47%) e formas recombinantes

(47%). O CRF02_AG foi a estirpe que apresentou maior prevalência em STP, conforme revelam os resultados, correspondendo a 31% dos infectados.

Embora S. Tomé e Príncipe apresente uma prevalência de infecção pelo VIH-1 considerada baixa (0,5%), este localiza-se próximo de uma região considerada o epicentro da epidemia (45), razão pela qual apresenta características epidemiológicas do VIH-1 semelhantes aos países vizinhos desta região da costa Ocidental, como os Camarões, o Gabão, e a Nigéria, caracterizados por uma alta variabilidade viral.

Este facto evidencia que o fenómeno migratório tem influência nas características epidemiológicas e na diversidade genética do VIH-1. Além disso, evidencia vulnerabilidades sociais que desta situação emergem, principalmente nos países de baixa e média renda, como São Tomé e Príncipe, onde as limitações económicas podem dificultar um controlo bem-sucedido da epidemia.

O fraco empenho na implementação de estudos epidemiológicos e na organização de estratégias públicas eficazes para o controlo da epidemia, constitui um entrave para o melhoramento de políticas de saúde.

O CRF02_AG é a variante do VIH-1 mais prevalente na África Ocidental e Central, onde representa cerca de 50% das infeções (49,53). A sua deteção noutras regiões do globo, como na Europa e América do Norte, tem sido cada vez mais comum. Este facto estará provavelmente relacionado com fluxos migratórios de migrantes provenientes destas regiões endémicas em CRF02_AG para essas regiões (49,53). Contudo, a causa do “sucesso” epidémico das diferentes linhagens do CRF02_AG noutras regiões é provavelmente mais complexa (54).

Por ser uma variante do VIH-1 predominante em África, um melhor conhecimento das suas características biológicas pode ter implicações importantes no diagnóstico das infeções por testes serológicos e moleculares e no tratamento (53).

Estudos conduzidos por Daiana et al. sugerem que a maior parte da diversidade atual do CRF02_AG vem da disseminação de algumas linhagens provenientes da África Central, introduzidas e disseminadas com sucesso na África Ocidental e Camarões entre o final da

década de 1960 e meados da década de 1980 (49). Estas formas recombinantes parecem ser o resultado do cruzamento entre os subtipos dentro de uma pessoa duplamente infetada, que posteriormente são passadas para outras pessoas, (55) e que se estabelecem com sucesso, formando uma nova linhagem.

O recombinante CRF02_AG deve ser o resultado da recombinação entre subtipos A e G, uma vez que este apresenta no seu genoma fragmentos dos dois subtipos, assim como descrito num estudo epidemiológico conduzido na Nigéria (27,56), tendo sido designado após o sequenciamento genómico quase completo de estirpes do tipo IbNg da Nigéria, Djibuti e Costa do Marfim (57). Contudo os estudos até então realizados não mostraram uma diferença significativa na sobrevida ou na progressão clínica da doença entre pessoas infetadas com o CRF02_AG, em comparação com aquelas infetadas com outros CRFs ou subtipos (53,55).

Não obstante o facto do CRF02_AG ser uma forma recombinante dos subtipos A e G, bastante prevalentes na África, todas as estirpes do subtipo G atualmente caracterizadas compartilham a mesma relação ambígua do subtipo A na sua região acessória, indicando uma possível ascendência recombinante dessa linhagem (27).

As variantes não-B do VIH-1 representam mais de 90% das estirpes circulantes em todo o mundo, com o subtipo C representando 50% de todas as infeções no mundo (58). Com o aumento da mobilidade e da migração, as variantes do VIH-1 inevitavelmente se misturam em diferentes partes do mundo, e a distribuição de diferentes formas de HIV-1 no mundo é, portanto, um processo dinâmico (58).

Existem evidências de que o processo de genotipagem de subtipos não-B apresenta ainda incongruências entre os vários algoritmos na interpretação dos resultados, possivelmente por sua otimização ter sido originalmente efetuada para os vírus do subtipo B (45), e de que existem diferenças relacionadas ao desempenho dos ensaios de diagnóstico e à eficiência dos medicamentos antirretrovirais entre as diversas variantes do HIV-1. Portanto, o acompanhamento da epidemiologia molecular do VIH-1 e a caracterização genética das suas estirpes prevalentes continua a ser importante (9).

Os resultados obtidos num estudo conduzido por *Henry Requejo et al.* na África Subsaariana sugeriram que os subtipos A, C, D e CRF01_AE do VIH-1 estão bem adaptados para transmissão heterossexual enquanto o subtipo B é menos eficientemente transmitido por esta via (9). Por outro lado, na América do Norte, Europa Ocidental, Sul da Ásia e Índia, o subtipo B é eficientemente transmitido por usuários de drogas intravenosas, e entre indivíduos homossexuais, nos quais a infeção pelos subtipos A, C, D e CRF01_AE do HIV-1 parece ocorrer menos frequentemente (9).

5.2 Filogenia e Filogeografia

O principal objetivo deste estudo era determinar a origem da epidemia do subtipo mais prevalente identificado em S. Tomé e Príncipe: o CRF02_AG.

O método Bayesiano de filogeografia implementado no software BEAST permitiu a identificação do ancestral comum do CRF02_AG, que mais provavelmente terá dado início à epidemia de VIH-1 CRF02_AG em S.Tomé e Príncipe. Neste estudo estimámos a data do início da epidemia como tendo ocorrido entre 1983 e 1986.

Para além disso, os resultados da análise da árvore MCC revelaram que o CRF02_AG existente em S.Tomé e Príncipe foi introduzido a partir de um ancestral comum mais recente proveniente do Gana. Este resultado pode ser explicado por fluxos migratórios entre S.Tomé e Príncipe e os países da costa Ocidental Africana.

Apesar de S.Tomé e Príncipe não ter uma ligação fronteiriça direta com o Gana, e ligações migratórias diretas com este país, é possível que através de fluxos migratórios entre o arquipélago e os países vizinhos como o Gabão, os Camarões e a Nigéria, tenham surgido casos de indivíduos santomenses que contraíram a infeção com estirpes VIH-1 CRF02_AG provenientes do Gana. Sabe-se que os países da Costa Ocidental Africana sempre tiveram uma grande relação de intercâmbio de pessoas e mercadorias, uma vez que partilham fronteiras com livre circulação.

Especula-se que no período após a independência tenha havido uma intensificação das relações entre as populações do arquipélago e as de alguns países limítrofes da costa ocidental Africana, tais como Angola, Camarões, Gabão, Guiné Equatorial e Nigéria (45). Neste

período foram observados movimentos populacionais dos são-tomenses para estes países, tendo também dado origem a um aumento considerável das ISTs (infecções sexualmente transmissíveis) (PNLS, 2003) (45).

Neste processo migratório muitos santomenses que haviam migrado começaram a retornar ao país à procura de melhores condições de acesso ao tratamento, correspondendo com o pico da epidemia no mundo. Além disso, houve um período (a partir dos anos 2001) em que se verificou um importante fluxo migratório de indivíduos provenientes da Nigéria para S.Tomé, coincidente com o período da especulação do petróleo no país. O que também pode ter contribuído para casos de infeção por CRF02_AG em S.Tomé, embora tivessem sido poucas as sequências santomenses que formaram clusters com estirpes CRF02_AG da Nigéria.

A infeção por VIH-1 em S.Tomé e Príncipe caracteriza-se por grande variabilidade genética, incluindo diversos subtipos e formas recombinantes, características estas que partilha com alguns países da costa Ocidental Africana. Os nossos resultados confirmam também que, para além de ter havido múltiplas introduções de subtipos virais no país, houve também múltiplas introduções das estirpes do CRF02_AG, que se distribuem em clusters diversos da árvore. Estes factos ressaltam a necessidade de mais investigações na área da biologia molecular do VIH-1 em S.Tomé, para aumentar a robustez dos estudos nesta área.

Sendo o VIH-1 CRF02_AG o subtipo mais prevalente no país (S.Tomé e Príncipe), associado à sua localização geografia, pode-se inferir a partir deste estudo que o início da epidemia tenha se dado com a introdução da estirpe CRF02_AG proveniente do Gana num período compreendido entre os anos de 1983 e 1986, uma vez que o CRF02_AG é o subtipo mais prevalente neste país (Gana) (57), facto que explicaria também a rápida dispersão dessas estirpes para países vizinhos da costa.

Os resultados deste estudo evidenciam uma vasta rede de disseminação eficiente do VIH-1 CRF02-AG para e entre os países da Costa Ocidental Africana, assim como para países do continente europeu, como Portugal, Espanha, Itália, Suíça e Suécia; além de países da América do Norte, como os Estados Unidos. Após estas introduções, redes de transmissão

complexas e autóctones de estirpes CRF02_AG terão ocorrido nestes países. Contudo, num estudo conduzido por *Edson Delatorre et al*, sugere que o facto da prevalência do CRF02_AG nestes dois continentes não ser mais elevada, se deva a sua fraca transmissibilidade em redes de transmissão por via sexual, que será a principal via de transmissão deste subtipo (54).

Por ser uma forma recombinante, a análise filogenética do CRF02_AG é um processo complexo, a até certo ponto controverso. Vários estudos demonstraram que a origem deste recombinante tenha ocorrido inicialmente na África central, com posterior introdução e disseminação na África Ocidental e nos Camarões, e dispersão para outras partes do mundo. No entanto, o padrão espaço-temporal de disseminação do CRF02_AG em escala global permanece em grande parte desconhecido (49), sugerindo ainda que a fraca informação obtida na análise filogeográfica do CRF02_AG proveniente da África Ocidental é o reflexo de múltiplas e frequentes trocas virais entre os países desta região, devido a forte mobilidade humana entre os mesmos (54).

Assim, utilizando-se ferramentas da biologia molecular pode-se deduzir a origem dos VIHs em diversos grupos populacionais, detetar e monitorizar introdução de novas variantes num país ou numa determinada região, contribuindo para melhor compreensão da filodinâmica deste vírus (45).

5.3 Resistência aos ARVs

O aumento do acesso ao tratamento do VIH-1 no mundo em desenvolvimento e o aumento da infecção pertencente ao subtipo não-B por meio de viagens e migrações colocam novas questões sobre a suscetibilidade e a resposta desses diversos vírus do VIH-1 a medicamentos antirretrovirais (59).

Os dados iniciais das coortes de tratamento na África levantaram duas preocupações: primeiro, que certos subtipos de VIH-1 podem se espalhar ou progredir mais rapidamente do que outros, tornando as decisões de tratamento mais desafiadoras; e que o tratamento voltado para o subtipo B pode não ser aplicável aos subtipos não-B (55).

A revisão de publicações que examinaram a resposta entre indivíduos infectados com estirpes não-B, tratados com antirretrovirais, demonstra que cerca de 60% destes indivíduos com as estirpes não-B que falham na terapia antirretroviral, têm evidências genotípicas de resistência aos medicamentos (59). O número de pessoas que recebem TARV em países com recursos limitados aumentou significativamente nos últimos anos, principalmente na África subsaariana e no sul e sudeste da Ásia, as duas áreas em que residem mais de 90% dos indivíduos necessitados de TARV (58).

Este estudo, realizado com 88 sequências do gene *pol* do VIH-1 em formato FASTA submetidas ao algoritmo da *Stanford University HIValg Program* (<https://hivdb.stanford.edu>), permitiu estimar uma prevalência que consideramos elevada (18%) de resistência aos antirretrovirais em S. Tomé e Príncipe, tendo em conta que a TARV foi introduzida no país apenas em 2004/2005 e as amostras foram colhidas de 2004 a 2007. Este facto é preocupante para a realidade do país, uma vez que na altura em que foram colhidas as amostras, 18% dos infectados já apresentavam mutações de resistência, e atualmente a situação poderá ser bastante mais grave. Existe grande probabilidade de haver um incremento da resistência, uma vez que não existem meios para se fazer a vigilância das resistências aos antirretrovirais, dado que o país não dispõe de recursos financeiros suficientes para aquisição de testes genotípicos de resistência aos ARVs. No entanto, realçamos que os dados analisados são antigos, pelo que não sabemos a realidade actual de resistência aos ARVs em STP.+

Os resultados revelaram ainda que as classes de fármacos com maior prevalência de mutações de resistência foram os NNRTIs (16%), seguida pelos NRTIs (15%). Estas duas classes de fármacos são as mais utilizadas no esquema de tratamento de primeira linha do VIH-1 no país, conforme as recomendações da OMS. Embora por volta dos anos 2016 e 2017 os esquemas de tratamento do VIH tenham mudado, com a introdução do esquema do combinado de três fármacos num só comprimido, algumas pessoas mantêm ainda os esquemas antigos por falta disponibilidade suficiente destas medicações para todos os infetados.

Outro aspeto positivo é que a resistência tripla (NRTI+ NNRTI+ PI) aos ARVs é baixa (1%), embora 9% (8/88) dos indivíduos da amostra tenham apresentado resistência dupla aos NRTI+ NNRTI, que são as classes de fármacos mais utilizadas no tratamento do VIH-1 em São Tomé e Príncipe. Este facto realça que o uso indiscriminado e não vigiado destes fármacos leva ao aumento das resistências.

Com estes dados, as autoridades de saúde e os coordenadores do programa nacional de luta contra a SIDA (PNLS) devem ficar em alerta para o tipo de tratamento que tem sido feito no país, dado que os testes de genotipagem não podem ser feitos antes do início da terapia antirretroviral por indisponibilidade financeira, associado a fraca monitorização dos pacientes que realmente estão em tratamento (taxa de cobertura da TARV em 2017 = 78,2%) e dos que abandonam o tratamento (taxa de abandono = 0,8%) (47). Por estes motivos, estes dados epidemiológicos podem nem sempre corresponder a realidade dos factos, uma vez que a fidedignidade dos mesmos é fraca, por aspetos relacionados com as fragilidades sociais e sanitárias do país.

As TARVs até hoje desenvolvidas são direcionadas para os estirpes do subtipo B, embora a resposta terapêutica entre os subtipos pareça não ser muito discrepante (55,59). Contudo, a emergência de mutações e polimorfismos encontrados em vários subtipos antes do início da terapia antirretroviral podem afetar as vias genéticas de resistência e a resposta à terapia antirretroviral (55). Uma vez que a utilização destes medicamentos tem aumentado nos países onde predominam os subtipos não-B e recombinantes, e onde os recursos e o acesso ao teste de genotipagem para a resistência aos ARVs são limitados, poderemos encontrar surpresas futuras relativas à emergência de resistências aos ARVs nestas estirpes (55).

O esquema da TARV foi introduzido em São Tomé e Príncipe em 2004 sob a tutela do Programa Nacional de Luta contra a SIDA (PNLS) (45). E os esquemas de tratamento que eram utilizados anteriormente, diferem em alguns aspetos dos esquemas atuais utilizados a partir do início de 2016.

Os fármacos de primeira linha que eram utilizados para o tratamento do VIH-1 em São Tomé e Príncipe no período de 2004 a 2007, época em que foram colhidas as amostras para o estudo correspondiam as seguintes combinações 2 NRTI e 1 NNRTI com os seguintes esquemas:

AZT+3TC (dose fixa combinada) + NVP (zidovudina 300 mg + lamivudina 150 mg) + nevirapina 200 mg); AZT+3TC (dose fixa combinada) + (zidovudina 150 mg + lamivudina 150 mg) + (efavirenz 600 mg); TDF+FTC (dose fixa combinada) (tenofovir 300 mg + emtricitabina 200 mg) + (efavirenz 600 mg).

Ao se analisar a percentagem dos níveis de resistência por fármacos, verifica-se que a Lamivudina (NRTI), fármaco de primeira linha apresenta 10% resistência de alto nível, assim como a Emtricitabina, também com 10% de resistência de alto nível. A Nevirapina, fármaco incluído na classe dos NNRTIs, também apresenta percentagem elevada de resistência de alto nível (10%), seguido pelo Efavirenz, com 8%. A Zidovudina, um fármaco bastante utilizado nos esquemas de tratamento do VIH-1 em S.T.P, tem valores baixos de resistência de alto nível, com 1%, e apresenta apenas 2% de resistência moderada. Já a Lamivudina, que é um fármaco bastante utilizado nos esquemas terapêuticos em São Tomé e Príncipe, tanto no tratamento de primeira linha, assim como de segunda linha, ou nos esquemas de pacientes co-infetados com o vírus da Hepatite B e/ou com o *Mycobacterium Tuberculosis*, apresenta 19% de resistência de alto nível (9/88). Estes dados podem ser preocupantes, uma vez que este é um dos principais fármacos usados nos esquemas de tratamento ARVs, e pode implicar uma resposta terapêutica pouco eficaz no tratamento dos infetados, propiciando riscos de aumento da resistência transmitida para os novos infetados.

A Emtricitabina, fármaco da classe dos NRTIs, que também apresenta valores relevantes na percentagem de resistência de níveis altos, põe também em causa a sua manutenção nos esquemas de tratamentos ARVs nos infetados no país. Uma das principais mutações de resistência (M184VI) encontrada neste fármaco (Emtricitabina) e na Lamivudina não é no entanto contra-indicação ao tratamento com as mesmas (60). O Abacavir, fármaco utilizado no esquema de segunda linha do tratamento ARV, apresenta uma prevalência de 10% de resistência moderada, com 1% dos indivíduos com resistência alta. Sendo assim, este fármaco pode ser mantido nos esquemas de tratamento, mas deve-se fazer atenção para o risco do aumento do número de casos de infetados com de resistência alta.

Uma importante limitação deste estudo é o facto das amostras terem sido colhidas no período de 2004 a 2007, podendo portanto não corresponder à realidade atual. Por isso, torna-se

importante a realização de um estudo com amostras mais recentes para reavaliar os padrões de resistência e mutações aos ARVs que existem no país no momento. Portanto, uma vigilância continuada da resistência transmitida (TDR) e da falha terapêutica é essencial em todos os subtipos e recombinantes do VIH (29).

As mutações de resistência associadas a fármacos NRTI identificados para Lamivudina são (M184V, T69D), para a Emtricitabina (M184V, T69D), Zidovudina (D67N, K70R, L210W, T215Y, M41L, K219Q), Abacavir (M184V, M41L, L210W, T215Y), Tenofovir (M41L, L210W, K70R, T69D) e Estavudina (T69D).

As mutações M184VI são mutações mais comuns associadas a resistência do vírus aos NRTIs, que embora causem resistência in vitro de alto nível a Lamivudina e Emtricitabina, não são contra-indicações a estes dois fármacos porque aumentam a suscetibilidade ao Tenofovir e a Zidovudina e diminuem a aptidão para replicação vira (45,60).

As inserções T69 são mutações que causam resistência a múltiplos fármacos e ocorrem com as TAMs (Mutações análogas da timidina). Já a mutação T69D reduz a suscetibilidade a Estavudina e Didanosina, que já não são recomendados para o tratamento do VIH (60).

A mutação D67E é uma mutação acessória dos TAMs (60), e é considerada uma mutação polimórfica (45).

As mutações D67N, K70R, L210W e M41L estão associadas a suscetibilidade reduzida ou baixa resposta virológica a Zidovudina (60). As mutações K70E/G ocorrem em pessoas tratadas com Tenofovir em regime combinado com Lamivudina ou Emtricitabina (45).

As mutações L210W, T215Y e M41L do tipo análogos da timidina (*TAMs*) estão associadas a suscetibilidade reduzida ao Abacavir e ao Tenofovir, quando em combinação com outras mutações de resistência a NRTI (60), podendo assim estar relacionadas com o insucesso terapêutico (45).

As mutações de resistência aos NNRTIs identificadas no estudo são: para a Nevirapina (K103N/S, V106M, Y181C, G190A), que são mutações *major* associadas a suscetibilidade

reduzida de alto nível do vírus a este fármaco; para a Rilpivirina, as mutações Y181C, G190A reduzem a suscetibilidade do vírus a este fármaco; e para o Efavirenz a mutação V106M está também associada a suscetibilidade reduzida de alto nível do vírus a este fármaco, e a mutação Y181C associada a redução da resposta virológica a este fármaco em combinação com outras mutações de resistência aos NNRTI. E por último, a mutação G190A está associada a suscetibilidade reduzida a Rilpivirina (60).

As mutações A98G e E138Q observadas em doentes tratados com Doravirina, Efavirenz, Etravirina, Nevirapina, Rilpivirina, são mutações não polimórficas associadas à resistência intermediária a alta à Rilpivirina (60). A mutação A98G, embora não polimórfica, ocorre em posições altamente polimórficas e a sua prevalência entre pacientes com NNRTI também é baixa (60).

A mutação V179D em associação com a mutação K103R reduz em dez vezes a suscetibilidade do vírus a Nevirapina e Efavirenz. As mutações Y181C e V179F em sinergia causam alto nível de resistência a Etravirina e Rilpivirina (60).

Há relatos acumulados da mutação K103N e outras importantes mutações relacionadas com resistência a medicamentos NNRTI em pessoas não tratadas (59). Além de que a K103N causa elevado nível de resistência à Nevirapina (NVP), Delavirdina (DLV), e Efavirenz (EFV) (45).

As mutações no local de ligação da RT reduzem a afinidade do vírus aos *NNRTIs* e levam ao insucesso terapêutico, nomeadamente A98G, K103N, K103S, V179E, Y181C e G190A (45). E em contraste com as reduções significativas na aptidão replicativa observadas com resistência a outras classes de fármacos, com os NNRTIs, alterações de um único nucleótido podem resultar em resistência de alto nível com apenas uma pequena perda de aptidão reprodutiva (33).

As mutações de resistência identificadas e associadas aos inibidores de protease foram as seguintes: Lopinavir (V32I, M46I, I47V, I54L, V82T, I84V), Atazanavir (V32I, M46I, I47V, I54L, V82T, I84V), Darunavir (V32I, I47V, I54L, I84V).

A mutação V32I reduz a resposta virológica a Darunavir e a Lopinavir, e apresenta suscetibilidade viral reduzida em combinação com outras mutações de resistência a PI (60). A mutação M46I também está associada a suscetibilidade reduzida do vírus ao Atazanavir e ao Lopinavir em combinação com outras mutações de resistência a PI (60). A mutação I47V representa uma resposta virológica reduzida às três drogas acima descritas em combinação com outras mutações de resistência aos inibidores de protease.

A mutação I54L representa suscetibilidade reduzida do vírus ao Darunavir e ao Lopinavir, e resposta virológica reduzida a Atazanavir em combinação com outras drogas. Já a mutação V82T causa uma resposta virológica reduzida a Lopinavir, e reduz a suscetibilidade viral a Atazanavir quando associada a uma outra mutação PI (60). A mutação I84V reduz a resposta virológica a Lopinavir e ao Darunavir, e uma suscetibilidade reduzida de alto nível ao Atazanavir (60).

A F53L é uma mutação de resistência acessória não polimórfica comum, e reduz a suscetibilidade do vírus a vários inibidores de protease, mesmo na ausência de outras mutações SDRMs (60).

Todas estas mutações aparecem associadas ao tratamento com os antirretrovirais, em pacientes expostos a Zidovudina (AZT), Lamivudina (3TC), Estavudina (D4T) e Nevirapina (NVP), utilizadas no esquema de tratamento dos infectados em S. Tomé e Príncipe (45).

Por isso, entre as várias possíveis explicações dadas para a alta prevalência de resistência em alguns países, incluem regimes de tratamento sub-ótimos, interrupção no fornecimento de medicamentos antirretrovirais e baixa adesão, levando à supressão e seleção inadequadas de resistência (59).

Um estudo conduzido por Martine Peeters *et al.* relata uma relação observada entre o subtipo genético e a resistência natural contra fármacos anti-retrovirais (58).

Neste estudo, identificamos que o subtipo B é o segundo subtipo com maior prevalência de mutações de resistência (60%) e a quinta estirpe viral mais prevalente (5%) nos infectados

da amostra, estando atrás do subtipo F, com 75% das mutações de resistência. Estas duas estirpes virais com maiores prevalências de mutações de resistência, implica que indivíduos infectados com as mesmas têm maiores probabilidades de apresentar mutações de resistência aos ARVs, comparativamente ao CRF02_AG, a estirpe mais prevalente, que apresenta 7% das mutações de resistência. Alguns estudos relatam que o subtipo B é o que geralmente apresenta mais mutações de resistência. A maioria das principais mutações de resistência encontradas no subtipo B também são encontradas nos subtipos não-B, embora a prevalência de algumas mutações seja mais alta em certos subtipos (59,60). No entanto, a resistência aos medicamentos não é bem estudada nos subtipos, CRFs e URFs que são menos comuns (59). A diversidade genética do VIH-1 existente S.Tomé e Príncipe pode ter implicações na escolha da terapia antirretroviral utilizada no país, uma vez que a maioria dos subtipos virais que circulam no mesmo são não-B e recombinantes. As taxas crescentes de transmissão do vírus resistente a medicamentos observadas nos últimos anos, juntamente com a pior resposta ao tratamento em pessoas com vírus resistente a medicamentos, têm que ser consideradas como uma base para a recomendação de que testes de resistência sejam realizados por rotina a pessoas com nova infecção pelo VIH (59).

A epidemia é globalizada e os subtipos não se limitam a uma só região, portanto é importante haver uma atenção redobrada aos subtipos não-B e CRFs. Além disso, a demanda por testes de resistência a medicamentos e carga viral para estirpes não-B aumentará significativamente, desta forma, é importante que esses testes tenham, tanto quanto possível, desempenho igual a todas as variantes do HIV-1 (55,58). E como os regimes contendo NNRTIs são a espinha dorsal do esquema de primeira linha preferencial atualmente recomendado pela OMS, e porque o teste de HIVDR não está prontamente disponível em países de baixa e média renda, as pessoas HIVDR em pré-tratamento podem não ser identificadas e, portanto, podem não receber regime eficaz promovendo assim, inadvertidamente, a transmissão do VIH mais resistentes aos medicamentos (34).

6. Conclusão

O VIH-1 é o retrovírus com maior importância epidemiológica a nível global, devido a sua dispersão mundial, as particularidades do seu genoma, sua vasta capacidade de mutação, replicação e adaptação ao humano, e taxas de mortalidade altas relacionadas com a SIDA (9).

Várias hipóteses e tentativas foram e têm sido feitas com o objetivo de se identificar a introdução do vírus na população humana. Análises epidemiológicas e filogenéticas atualmente disponíveis implicam que o VIH foi introduzido na população humana por volta de 1920 a 1940 (2).

Para que o planeamento em saúde pública se torne efetivo e bem-sucedido, precisamos de compreender a dinâmica das epidemias regionais e intervir apropriadamente. As ferramentas da epidemiologia molecular do VIH, implementadas em análises filogenéticas, filodinâmicas e filogeográficas, provaram ser ferramentas poderosas para o fornecimento de dados sobre a dinâmica da transmissão, parâmetros epidemiológicos críticos (número efetivo de infeções, tempos de geração, tempo entre a infeção e o diagnóstico) ou características espaço-temporais das epidemias (61).

A dispersão epidêmica do HIV-1 grupo M a partir de Kinshasa, capital da República Democrática do Congo (RDC), desde a primeira parte do século XX, resultou na extensa diversidade de subtipos, sub-subtipos, circulantes formas recombinantes (CRFs) e formas recombinantes únicas (URFs) relatadas em todo o mundo (49). Acredita-se que a pandemia do grupo M do VIH-1 provavelmente tenha sido desencadeada em parte por profundas alterações socio-comportamentais e industriais na África pós-colonial (62).

A eliminação do VIH como ameaça a saúde pública exige esforços acelerados para expandir a cobertura e a qualidade do tratamento dos TARVs e Serviços. Esta expansão precisa ser balanceada através de esforços bem coordenados para garantir que permaneça eficaz e que o risco e o potencial impacto das resistências aos antirretrovirais sejam minimizados (34).

A semelhança dos países da costa ocidental africana a epidemia do VIH-1 em S.Tomé e Príncipe caracteriza-se por uma diversidade genética, em que o CRF02_AG é o subtipo mais prevalente. Neste estudo, foi possível determinar pela primeira vez a sua origem da sua introdução no arquipélago.

Por não existirem estudos desta natureza, relacionados com a introdução dos VIHs no arquipélago, este trabalho de investigação permitiu através dos préstimos da filogenia molecular determinar a origem das estirpes CRF02_AG que circulam em S.Tomé, e a provável causa do início da epidemia do VIH-1 no país em meados dos anos 80, correspondendo com o período da pós-independência, aquando da entrada dos militares de vários países com os quais o STP mantém intercâmbios (45), e à intensificação das relações entre as populações do arquipélago e as de alguns países limítrofes da costa ocidental Africana, tais como Angola, Camarões, Gabão, Guiné Equatorial e Nigéria.

Embora a descoberta da epidemia da SIDA no mundo tenha sido identificada em 1981, com a descoberta do VIH em 1983, os primeiros casos da doença começaram a surgir em finais dos anos 70 nos Estados Unidos, chamando a atenção dos agentes de saúde pública neste país e nos países da Europa, entre outros. Entretanto, a doença já se encontrava instalada no continente africano há muito tempo, local da provável origem e início da epidemia. Isto propiciou o avanço da doença para vários países do globo, associado aos movimentos migratórios em massa, a pobreza, ao uso de drogas injetáveis, entre outros fatores.

Vários estudos foram e têm sido feitos na tentativa de verificar se existem diferenças nas taxas de progressão da doença entre os diferentes subtipos. Sendo que alguns desses estudos não mostraram haver diferenças significativas na sobrevida ou na progressão clínica da doença entre pessoas infetadas com CRF02_AG, em comparação com aquelas infetadas com outros CRFs ou subtipos (55). Mas sabe-se que interações entre o hospedeiro e o VIH-1 variam de acordo com o subtipo (55) e que os fatores que levam a esses padrões são complexos e multifatoriais e incluem possíveis diferenças biológicas entre os subtipos (29). Além disso, a progressão da doença pode estar relacionada com fatores como o acesso a cuidados médicos, estado nutricional, fatores genéticos do hospedeiro, redes de transmissão e modo de transmissão viral (por exemplo, sexual, drogas injetáveis ou vertical) (29,55).

Com a descoberta da terapia antirretroviral muitas vidas puderam ser salvas, embora ainda não se tenha conseguido controlar a progressão da epidemia em escala mundial, os progressos alcançados foram muitos. Contudo, o objetivo da Agenda 2030 das Nações Unidas para os Objetivos do Desenvolvimento Sustentável e Global, da OMS junto com a UNAIDS é acabar com a epidemia da SIDA até 2030, com a adoção da meta 90-90-90 (34).

O uso de esquemas específicos de medicamentos em diferentes regiões e populações e o desenvolvimento de estratégias de tratamento dependem da investigação e compreensão da resistência aos medicamentos no contexto de diversos subtipos virais (59). Muitas sequências com dados sobre doses, duração, adesão e resultados clínicos são necessárias para analisar esses padrões complexos de resistência a medicamentos em indivíduos infectados por subtipos não-B (51,59).

S.Tomé e Príncipe apresentou resultados considerados satisfatórios em relação ao perfil das resistências as drogas antirretrovirais, com uma prevalência até certo ponto baixa para a realidade do país, mas deve procurar formas de melhorar as estratégias para lidar com a doença. Pelo facto de ser um país pobre e com poucos recursos, fica vulnerável aos efeitos sociais negativos que advêm desta.

Por isso, as recomendações aqui feitas vão no sentido de alertar as autoridades sanitárias para o incremento de esforços que permitam uma melhor coordenação de estratégias de rastreio e monitorização dos infectados, com o intuito de manter ou diminuir o avanço da epidemia no país. É necessário por isso criar medidas de incentivo à realização de mais estudos ou trabalhos de investigação científica que promovam o melhoramento dos dados estatísticos e epidemiológicos, para que se possa trabalhar ou projetar a partir destes. É crucial acelerar os esforços para a realização de um estudo nacional de genotipagem das mutações de resistência, a fim de se conhecer o perfil de resistência das variantes virais VIH-1 existentes no país, assim como promover o rastreio e a monitorização desses padrões de resistência. Promover estratégias de reeducação sexual da população para a prática de relações sexuais protegidas, enaltecendo as vantagens uso de preservativos, e os riscos associados a aquisição

das ISTs (infecções sexualmente transmissíveis), com destaque para o VIH; é também importantíssimo para a prevenção da epidemia.

É finalmente necessário melhorar a qualidade do tratamento do VIH e aproximá-lo dos ideais universais de cobertura de saúde de serviços integrados, criando abordagens de cuidados de saúde centradas na comunidade e comunitárias, e responsabilidade compartilhada pela entrega eficaz de programas (31).

Este estudo procura ser um ponto inicial para outros estudos futuros que permitirão orientar e dirigir as estratégias acima referidas.

7. Referências Bibliográficas

1. Gallo, R. C., Salahuddin, S.Z., Popovic, M. Frequent Detection of Isolation of Cytopathic Retroviruses (HTLV-III) from patients with AIDS and Risk for AIDS. *Science*. 1984;224:500-503.
2. German Advisory Committee Blood (Arbeitskreis Blut), Subgroup 'Assessment of Pathogens Transmissible by Blood'. *Human Immunodeficiency Virus (HIV). Transfusion Medicine and Hemotherapy*. 2016;43(3):203–22.
3. Ganser-Pornillos BK, Yeager M, Pornillos O. Assembly and Architecture of HIV. Em: Rossmann MG, Rao VB, editores. *Viral Molecular Machines* [Internet]. Boston, MA: Springer US; 2012 [citado 10 de Maio de 2019]. p. 441–65. Disponível em: http://link.springer.com/10.1007/978-1-4614-0980-9_20
4. MMWR. HIV and AIDS in the United States 1981-2000. CDC. Junho de 2001;50:430.
5. Braunwald E, Harrison TR, editores. *Harrison's principles of internal medicine*. 15. ed. New York: McGraw-Hill; 2001.
6. HIV/AIDS [Internet]. [citado 26 de Junho de 2019]. Disponível em: <https://www.who.int/news-room/fact-sheets/detail/hiv-aids>
7. Salvador E. DOMINICAN REPUBLIC ECUADOR. :376.
8. Haile-Selassie H, World Health Organization, World Health Organization, Department of HIV/AIDS, Centers for Disease Control and Prevention (U.S.), Global Fund to Fight AIDS T and Malaria, et al. WHO HIV drug resistance report 2017. [Internet]. 2017 [citado 28 de Junho de 2019]. Disponível em: <http://apps.who.int/iris/bitstream/10665/255896/1/9789241512831-eng.pdf>
9. Requejo HIZ. Worldwide molecular epidemiology of HIV. *Rev Saude Publica*. Abril de 2006;40(2):331–45.
10. Fanales-Belasio E, Raimondo M, Suligoj B, Buttò S. HIV virology and pathogenetic mechanisms of infection: a brief overview. *Annali dell'Istituto Superiore di Sanità* [Internet]. Março de 2010 [citado 28 de Maio de 2019];(1). Disponível em: http://www.iss.it/binary/publ/cont/ANN_10_01_02.pdf
11. Genes and Genome of HIV-1. *Journal of Phylogenetics & Evolutionary Biology* [Internet]. 2014 [citado 22 de Maio de 2019];02(01). Disponível em:

<http://esciencecentral.org/journals/genes-and-genome-of-hiv-2329-9002-2-126.php?aid=27166>

12. Li G, Piampongsant S, Faria NR, Voet A, Pineda-Peña A-C, Khouri R, et al. An integrated map of HIV genome-wide variation from a population perspective. *Retrovirology* [Internet]. Dezembro de 2015 [citado 22 de Maio de 2019];12(1). Disponível em: <https://retrovirology.biomedcentral.com/articles/10.1186/s12977-015-0148-6>
13. Gao F, Bailes E, Robertson DL, Chen Y, Rodenburg CM, Michael SF, et al. Origin of HIV-1 in the chimpanzee *Pan troglodytes troglodytes*. *Nature*. 1 de Fevereiro de 1999;397(6718):436–41.
14. Li G, De Clercq E. HIV Genome-Wide Protein Associations: a Review of 30 Years of Research. *Microbiology and Molecular Biology Reviews*. Setembro de 2016;80(3):679–731.
15. Levy JA. HIV pathogenesis and long-term survival: AIDS. Novembro de 1993;7(11):1401–10.
16. Peçanha EP, Antunes OAC, Tanuri A. Estratégias farmacológicas para a terapia anti-AIDS. *Química Nova*. Dezembro de 2002;25(6b):1108–16.
17. Chan DC, Kim PS. HIV Entry and Its Inhibition. *Cell*. Maio de 1998;93(5):681–4.
18. Engelman A, Cherepanov P. The structural biology of HIV-1: mechanistic and therapeutic insights. *Nature Reviews Microbiology*. Abril de 2012;10(4):279–90.
19. Pollard VW, Malim MH. THE HIV-1 REV PROTEIN. *Annual Review of Microbiology*. Outubro de 1998;52(1):491–532.
20. Ferguson MR, Rojo DR, von Lindern JJ, O’Brien WA. HIV-1 replication cycle. *Clinics in Laboratory Medicine*. Setembro de 2002;22(3):611–35.
21. Chin BS. Molecular Epidemiology of Human Immunodeficiency Virus. *Infection & Chemotherapy*. 2017;49(1):1.
22. Simon V, Ho DD, Abdool Karim Q. HIV/AIDS epidemiology, pathogenesis, prevention, and treatment. *The Lancet*. Agosto de 2006;368(9534):489–504.

23. Faria NR, Rambaut A, Suchard MA, Baele G, Bedford T, Ward MJ, et al. The early spread and epidemic ignition of HIV-1 in human populations. *Science*. 3 de Outubro de 2014;346(6205):56–61.
24. Worobey M, Gemmel M, Teuwen DE, Haselkorn T, Kunstman K, Bunce M, et al. Direct evidence of extensive diversity of HIV-1 in Kinshasa by 1960. *Nature*. Outubro de 2008;455(7213):661–4.
25. Barre-Sinoussi F, Chermann J, Rey F, Nugeyre M, Chamaret S, Gruest J, et al. Isolation of a T-lymphotropic retrovirus from a patient at risk for acquired immune deficiency syndrome (AIDS). *Science*. 20 de Maio de 1983;220(4599):868–71.
26. McCloskey RM, Liang RH, Harrigan PR, Brumme ZL, Poon AFY. An Evaluation of Phylogenetic Methods for Reconstructing Transmitted HIV Variants using Longitudinal Clonal HIV Sequence Data. *Journal of Virology*. 1 de Junho de 2014;88(11):6181–94.
27. Robertson DL. HIV-1 Nomenclature Proposal. *Science*. 7 de Abril de 2000;288(5463):55d–5.
28. Yamaguchi J, Vallari A, McArthur C, Sthreshley L, Cloherty GA, Berg MG, et al. Brief Report: Complete Genome Sequence of CG-0018a-01 Establishes HIV-1 Subtype L. *JAIDS Journal of Acquired Immune Deficiency Syndromes*. Março de 2020;83(3):319–22.
29. Hemelaar J, Elangovan R, Yun J, Dickson-Tetteh L, Fleminger I, Kirtley S, et al. Global and regional molecular epidemiology of HIV-1, 1990–2015: a systematic review, global survey, and trend analysis. *The Lancet Infectious Diseases*. Fevereiro de 2019;19(2):143–55.
30. HIV Circulating Recombinant Forms (CRFs) [Internet]. [citado 22 de Setembro de 2019]. Disponível em: <https://www.hiv.lanl.gov/content/sequence/HIV/CRFs/CRFs.html>
31. World Health Organization. Consolidated guidelines on the use of antiretroviral drugs for treating and preventing HIV infection: recommendations for a public health approach. [Internet]. 2016 [citado 30 de Junho de 2019]. Disponível em: http://apps.who.int/iris/bitstream/10665/208825/1/9789241549684_eng.pdf
32. Günthard HF, Aberg JA, Eron JJ, Hoy JF, Telenti A, Benson CA, et al. Antiretroviral Treatment of Adult HIV Infection: 2014 Recommendations of the International Antiviral Society–USA Panel. *JAMA*. 23 de Julho de 2014;312(4):410.

33. Arts EJ, Hazuda DJ. HIV-1 Antiretroviral Drug Therapy. Cold Spring Harbor Perspectives in Medicine. 1 de Abril de 2012;2(4):a007161–a007161.
34. World Health Organization. GUIDELINES ON JULY 2017 THE PUBLIC HEALTH RESPONSE TO PRETREATMENT HIV DRUG RESISTANCE [Internet]. Geneva; 2017. Disponível em: <https://apps.who.int/iris/bitstream/handle/10665/255880/9789241550055-eng.pdf>
35. HIV Infection and AIDS Medication: Antiretroviral agent, nucleoside reverse-transcriptase inhibitor, Antiretroviral Agent, Protease Inhibitor, Antiretroviral agent, non-nucleoside reverse-transcriptase inhibitor, Antiretroviral Agent, Integrase Inhibitor, Antiretroviral Agent, Fusion Inhibitor, Antiretroviral agent, CCR5 antagonist, HIV, Entry Inhibitors, Complete Regimen Combinations, Antiretroviral Combinations, CYP3A4 Inhibitors, Antibiotic, Sulfonamide Derivative, Growth hormone releasing factor [Internet]. [citado 11 de Julho de 2019]. Disponível em: <https://emedicine.medscape.com/article/211316-medication>
36. Clavel F, Hance AJ. HIV Drug Resistance. New England Journal of Medicine. 4 de Março de 2004;350(10):1023–35.
37. Taiwo B. Understanding transmitted HIV resistance through the experience in the USA. International Journal of Infectious Diseases. Setembro de 2009;13(5):552–9.
38. Instituto Nacional de Estatística [Internet]. Finance. [citado 22 de Setembro de 2019]. Disponível em: <https://www.ine.st/>
39. Instituto Nacional de Estatísticas. IV Recenseamento Geral da População e Habitação – 2012: Características Económicas da População. S.Tomé, S.Tomé e Príncipe. Instituto Nacional de Estatísticas; 2014.
40. Ministerio da Saúde. Contas de Saúde 2014 e 2015. Ministério da Saúde; 2018 Setembro de.
41. PNUD. Relatório Nacional de Desenvolvimento Humano em São Tomé e príncipe. S.Tomé- A qualidade de liderança como factor inibidor da instabilidade política e promotor do desenvolvimento humano: o papel da Sociedade Civil e da Juventude. PNUD; 2014.
42. World Development Indicators | DataBank [Internet]. [citado 22 de Setembro de 2019]. Disponível em: <https://databank.worldbank.org/reports.aspx?source=World-Development-Indicators>

43. | Human Development Reports [Internet]. [citado 22 de Setembro de 2019]. Disponível em: <http://hdr.undp.org/en/2018-update>
44. PNLS. Plano Estratégico Multisectorial de Luta contra o HIV/SIDA: 2013-2017. Ministério da Saúde; 2013.
45. Bonfim I. HIV-1 na Republica Democrática de São Tomé e Príncipe: estudos sobre a situação atual. Instituto de Higiene e Medicina Tropical, Universidade Nova de Lisboa; 2008.
46. UNPD, INE-STP, Unicef. São Tome and Principe: Multiple Indicator cluster Survey 2014; Final Report. São Tome and Principe: Ministério da Saúde: Centro Nacional de Endemias; 2016.
47. PNLS. Resultados dos Dados de Monitorização e Avaliação do HIV/SIDA 2007 a 2017. S.Tomé e Príncipe: Ministério da Saúde: Centro Nacional de Endemias; 2018.
48. HIV Country Profiles [Internet]. [citado 23 de Setembro de 2019]. Disponível em: <http://cfs.hivci.org/country-factsheet.html#>)
49. Mir D, Jung M, Delatorre E, Vidal N, Peeters M, Bello G. Phylodynamics of the major HIV-1 CRF02_AG African lineages and its global dissemination. *Infection, Genetics and Evolution*. Dezembro de 2016;46:190–9.
50. Stanford. HIVDB Genotypic Resistance Test (GRT) Interpretation System [Internet]. Disponível em: <https://hivdb.stanford.edu/assets/media/genotypic-resistance-test-interpretation-system-oct2018.57d61710.pdf>
51. Abecasis AB, Lemey P, Vidal N, de Oliveira T, Peeters M, Camacho R, et al. Recombination Confounds the Early Evolutionary History of Human Immunodeficiency Virus Type 1: Subtype G Is a Circulating Recombinant Form. *Journal of Virology*. 15 de Agosto de 2007;81(16):8543–51.
52. Bonfim I, Carvalho AP, Carvalho AV, Sousa B, Palma C, Cabanas J, et al. HIV-1 molecular epidemiology in Sao Tome e Principe. *Wiley*; 2007:0501.
53. Vergne L, Bourgeois A, Mpoudi-Ngole E, Mougnotou R, Mbuagbaw J, Liegeois F, et al. Biological and genetic characteristics of HIV infections in Cameroon reveals dual group M and O infections and a correlation between SI-inducing phenotype of the predominant CRF02_AG variant and disease stage. *Virology*. Junho de 2003;310(2):254–66.

54. Delatorre E, Velasco-De-Castro CA, Pilotto JH, Couto-Fernandez JC, Bello G, Morgado MG. Short Communication: Reassessing the Origin of the HIV-1 CRF02_AG Lineages Circulating in Brazil. *AIDS Research and Human Retroviruses*. Dezembro de 2015;31(12):1230–1237.
55. Taylor BS, Sobieszczyk ME, McCutchan FE, Hammer SM. The Challenge of HIV-1 Subtype Diversity. *New England Journal of Medicine*. 10 de Abril de 2008;358(15):1590–602.
56. Tebit DM, Zekeng L, Kaptué L, Kräusslich H-G, Herchenröder O. Construction and characterisation of a full-length infectious molecular clone from a fast replicating, X4-tropic HIV-1 CRF02_AG primary isolate. *Virology*. Setembro de 2003;313(2):645–52.
57. Fischetti L, Opare-Sem O, Candotti D, Sarkodie F, Lee H, Allain JP. Molecular epidemiology of HIV in Ghana: Dominance of CRF02_AG. *Journal of Medical Virology*. Junho de 2004;73(2):158–66.
58. Peeters M, Esu-Williams E, Vergne L, Montavon C, Mulanga-Kabeya C, Harry T, et al. Predominance of Subtype A and G HIV Type 1 in Nigeria, with Geographical Differences in Their Distribution. *AIDS Research and Human Retroviruses*. Março de 2000;16(4):315–25.
59. Kantor R, Katzenstein D. Drug resistance in non-subtype B HIV-1. *Journal of Clinical Virology*. Março de 2004;29(3):152–9.
60. Tzou PL, Rhee S-Y, Descamps D, Clutter DS, Hare B, Mor O, et al. Integrase strand transfer inhibitor (INSTI)-resistance mutations for the surveillance of transmitted HIV-1 drug resistance. *Journal of Antimicrobial Chemotherapy* [Internet]. 16 de Outubro de 2019 [citado 20 de Outubro de 2019]; Disponível em: <https://academic.oup.com/jac/advance-article/doi/10.1093/jac/dkz417/5588580>
61. Paraskevis D, Nikolopoulos GK, Magiorkinis G, Hodges-Mameletzis I, Hatzakis A. The application of HIV molecular epidemiology to public health. *Infection, Genetics and Evolution*. Dezembro de 2016;46:159–68.
62. Abecasis A, Paraskevis D, Epalanga M, Fonseca M, Burity F, Bartolomeu J, et al. HIV-1 genetic variants circulation in the North of Angola. *Infection, Genetics and Evolution*. Abril de 2005;5(3):231–7.

63. Jacobs G. HIV-1 resistance analyses from therapy-naïve patients in South Africa, Tanzania and the characterization of a new HIV-1 subtype C proviral molecular clone. 2012.